Software Freedom
Law Center

Fall Conference
at Columbia Law School

CLE materials

November 2nd, 2018

Information Regarding New York CLE Credits:

Columbia Law School has been certified by the New York State Continuing Legal Education (CLE) Board as an Accredited Provider of CLE programs. Under New York State CLE regulations, this live non-transitional CLE Program will provide 7 credit hours that can be applied toward the Areas of Professional Practice requirement. This CLE credit is awarded only to New York attorneys for full attendance of the Program in its entirety. Attorneys attending only part of the program are not eligible for partial credit. Attendance is determined by an attorney's sign-in and sign-out, as shown in the Conference registers. On final sign-out, attorneys should also submit their completed Evaluation Form, provided at the Conference. Please note the NYS Certificates of Attendance will be sent to the email address as it appears in the register unless otherwise noted there.

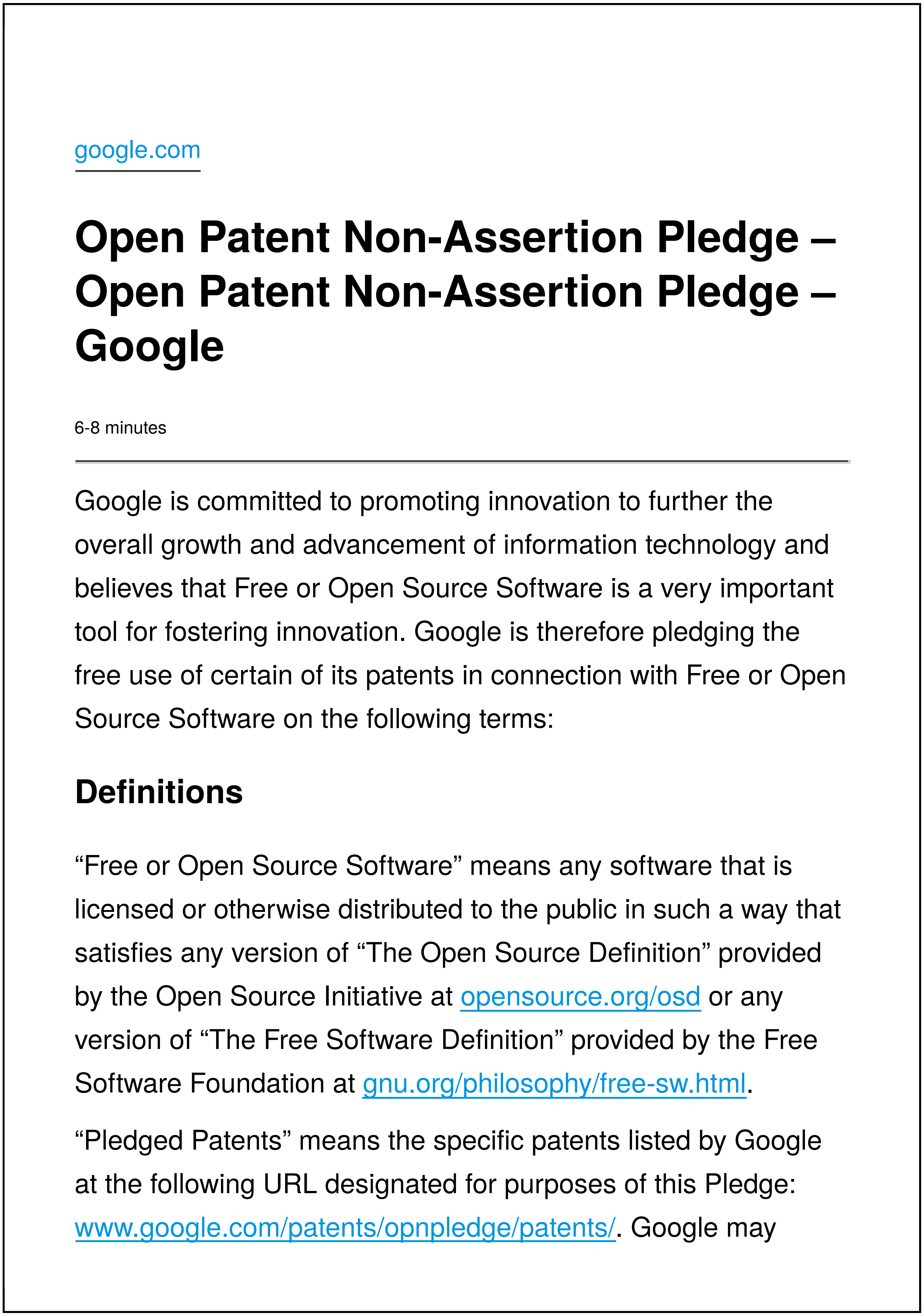
Contents

Part I

Google's Philosophy of FOSS

Chapter 1

Google Open Patent Non-Assertion Pledge



google.com

Open Patent Non-Assertion Pledge – Open Patent Non-Assertion Pledge – Google

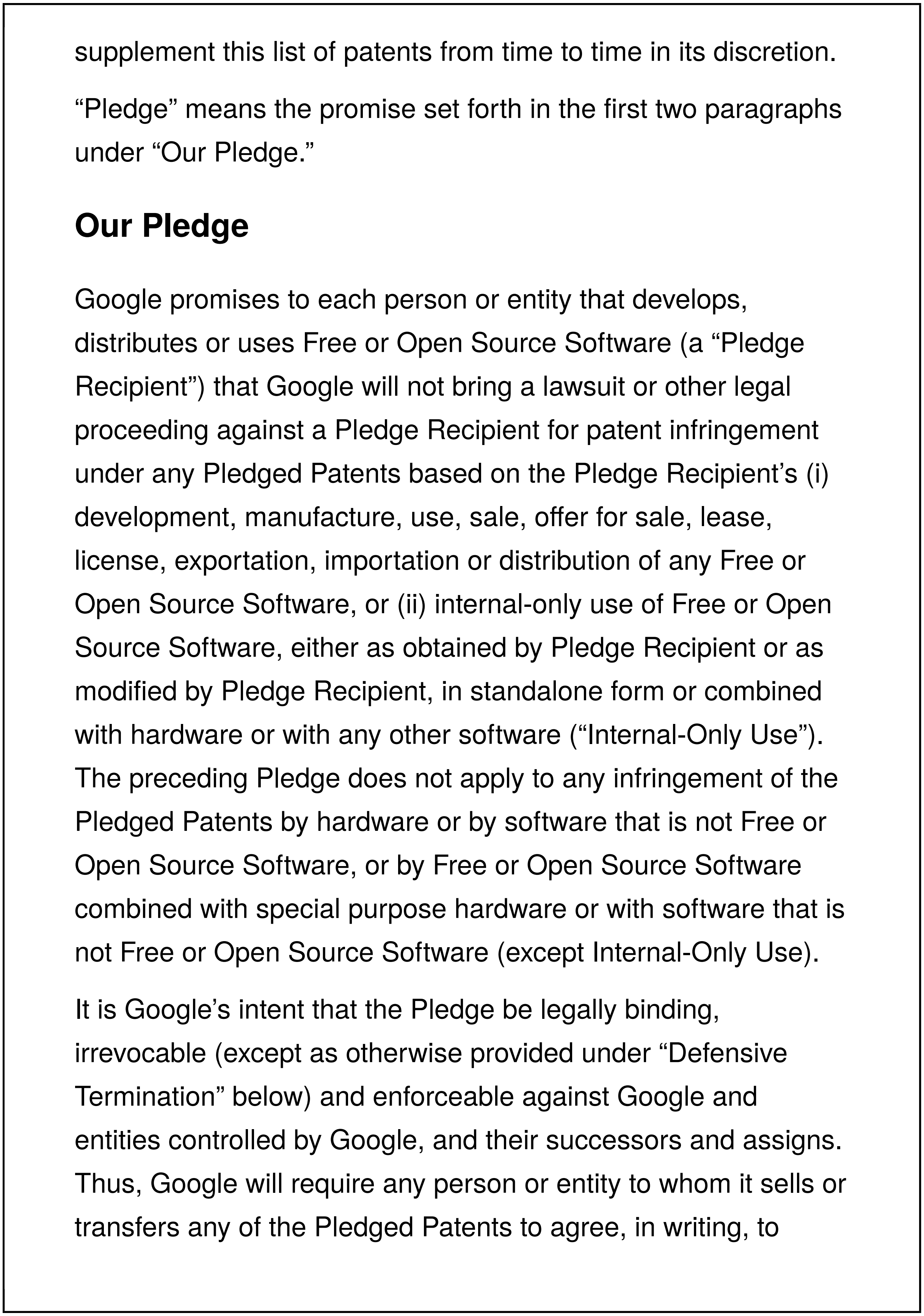
6-8 minutes

Google is committed to promoting innovation to further the overall growth and advancement of information technology and believes that Free or Open Source Software is a very important tool for fostering innovation. Google is therefore pledging the free use of certain of its patents in connection with Free or Open Source Software on the following terms:

Definitions

“Free or Open Source Software” means any software that is licensed or otherwise distributed to the public in such a way that satisfies any version of “The Open Source Definition” provided by the Open Source Initiative at opensource.org/osd or any version of “The Free Software Definition” provided by the Free Software Foundation at gnu.org/philosophy/free-sw.html.

“Pledged Patents” means the specific patents listed by Google at the following URL designated for purposes of this Pledge: www.google.com/patents/opnpledge/patents/. Google may



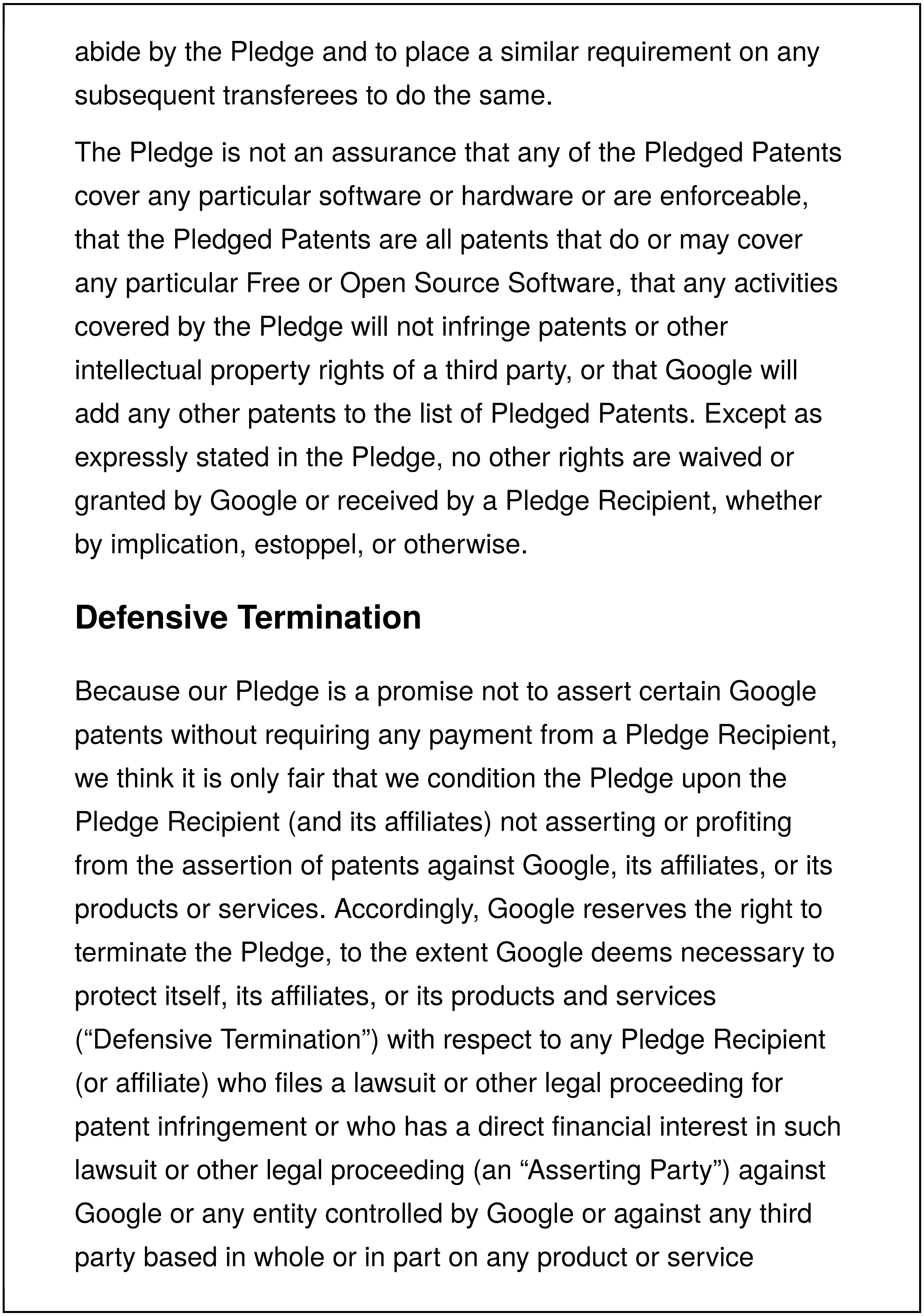
supplement this list of patents from time to time in its discretion.

“Pledge” means the promise set forth in the first two paragraphs under “Our Pledge.”

Our Pledge

Google promises to each person or entity that develops, distributes or uses Free or Open Source Software (a “Pledge Recipient”) that Google will not bring a lawsuit or other legal proceeding against a Pledge Recipient for patent infringement under any Pledged Patents based on the Pledge Recipient’s (i) development, manufacture, use, sale, offer for sale, lease, license, exportation, importation or distribution of any Free or Open Source Software, or (ii) internal-only use of Free or Open Source Software, either as obtained by Pledge Recipient or as modified by Pledge Recipient, in standalone form or combined with hardware or with any other software (“Internal-Only Use”). The preceding Pledge does not apply to any infringement of the Pledged Patents by hardware or by software that is not Free or Open Source Software, or by Free or Open Source Software combined with special purpose hardware or with software that is not Free or Open Source Software (except Internal-Only Use).

It is Google’s intent that the Pledge be legally binding, irrevocable (except as otherwise provided under “Defensive Termination” below) and enforceable against Google and entities controlled by Google, and their successors and assigns. Thus, Google will require any person or entity to whom it sells or transfers any of the Pledged Patents to agree, in writing, to



abide by the Pledge and to place a similar requirement on any subsequent transferees to do the same.

The Pledge is not an assurance that any of the Pledged Patents cover any particular software or hardware or are enforceable, that the Pledged Patents are all patents that do or may cover any particular Free or Open Source Software, that any activities covered by the Pledge will not infringe patents or other intellectual property rights of a third party, or that Google will add any other patents to the list of Pledged Patents. Except as expressly stated in the Pledge, no other rights are waived or granted by Google or received by a Pledge Recipient, whether by implication, estoppel, or otherwise.

Defensive Termination

Because our Pledge is a promise not to assert certain Google patents without requiring any payment from a Pledge Recipient, we think it is only fair that we condition the Pledge upon the Pledge Recipient (and its affiliates) not asserting or profiting from the assertion of patents against Google, its affiliates, or its products or services. Accordingly, Google reserves the right to terminate the Pledge, to the extent Google deems necessary to protect itself, its affiliates, or its products and services ("Defensive Termination") with respect to any Pledge Recipient (or affiliate) who files a lawsuit or other legal proceeding for patent infringement or who has a direct financial interest in such lawsuit or other legal proceeding (an "Asserting Party") against Google or any entity controlled by Google or against any third party based in whole or in part on any product or service developed by or on behalf of Google or any entity controlled by Google.

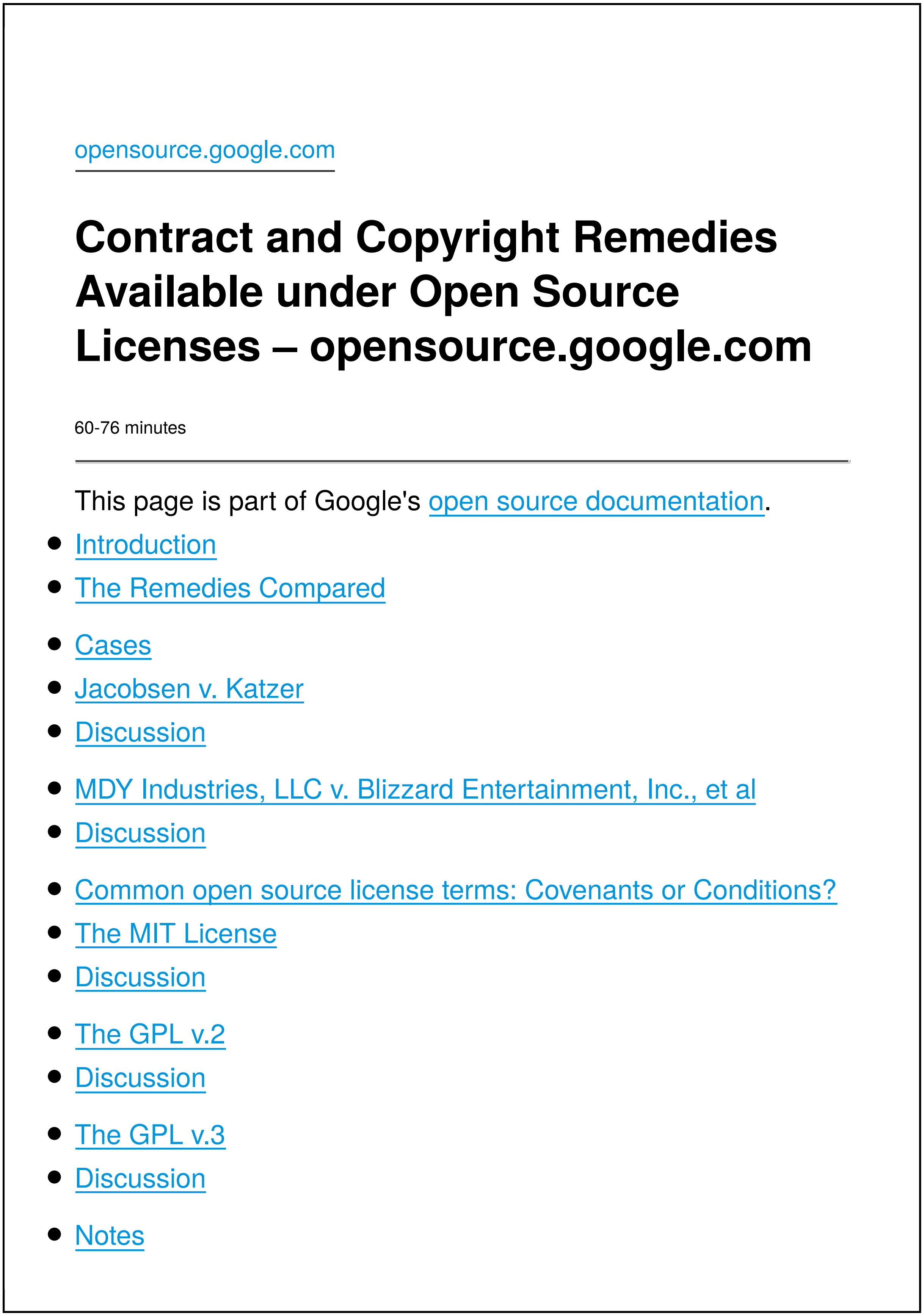
Any Defensive Termination by Google with respect to an Asserting Party shall have the same effect as if Our Pledge was never extended to such Asserting Party in the first instance. Google, in its sole discretion, shall determine the manner and terms, if any, by which rights under Pledged Patents may be extended to an Asserting Party after that Asserting Party's lawsuit or other legal proceeding is permanently dismissed, terminated or withdrawn in writing.

Mistaken Assertion

Should Google ever initiate a lawsuit or other legal proceeding for patent infringement based on software which is not the subject of a Defensive Termination, and then receive written notice from the party against whom such lawsuit or other legal proceeding has been filed providing sufficient information for Google to reasonably determine that such software in fact satisfies the requirements of the Pledge, then Google will use reasonable efforts to withdraw such lawsuit (or the applicable claims therein) or move to terminate such other legal proceeding (or the applicable portions thereof) within sixty (60) days after receiving such written notice.

Chapter 2

Contract and Copyright Remedies Available Under Open Source Licences

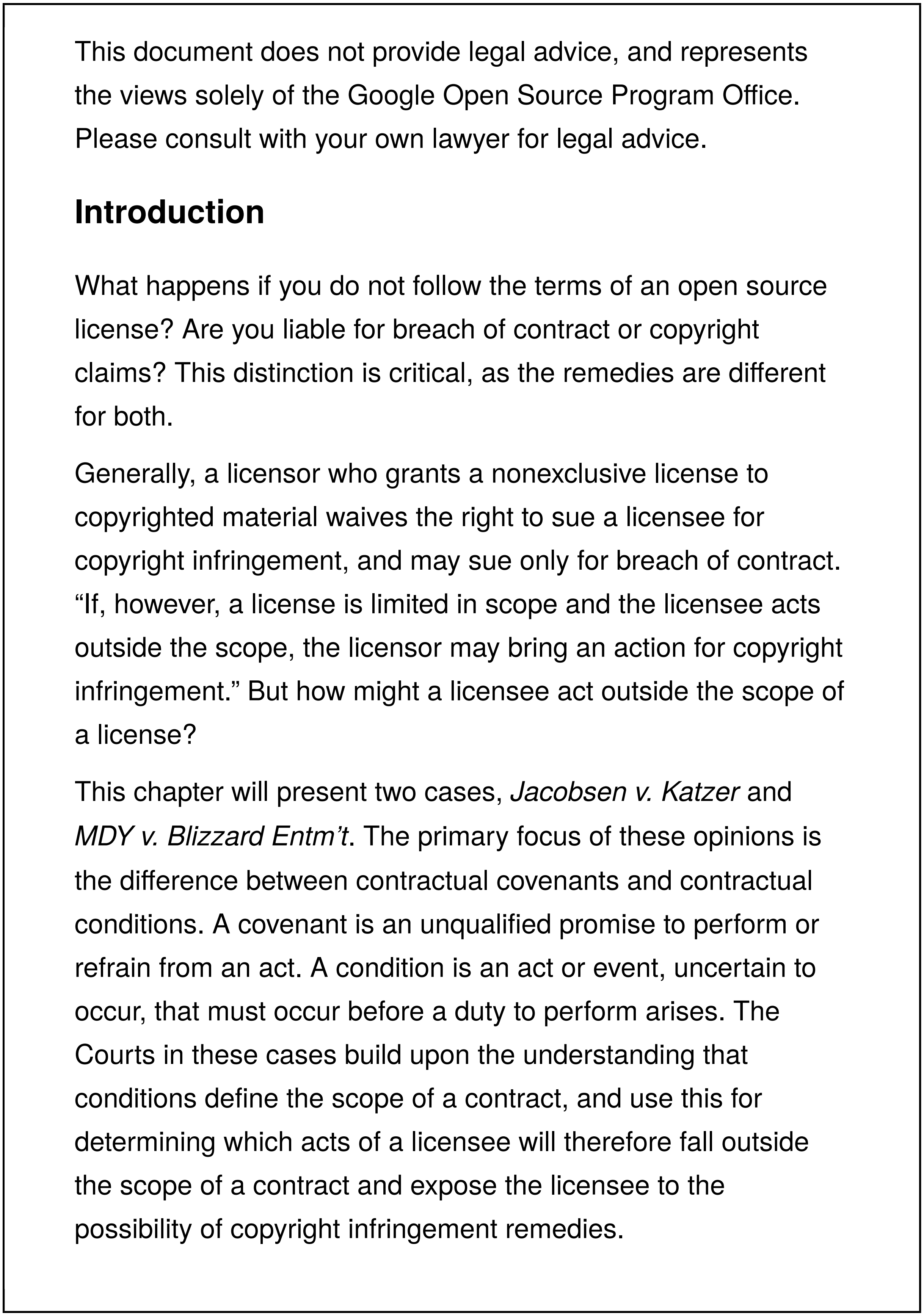


opensource.google.com

Contract and Copyright Remedies Available under Open Source Licenses – opensource.google.com

60-76 minutes

This page is part of Google's open source documentation.



This document does not provide legal advice, and represents the views solely of the Google Open Source Program Office. Please consult with your own lawyer for legal advice.

Introduction

What happens if you do not follow the terms of an open source license? Are you liable for breach of contract or copyright claims? This distinction is critical, as the remedies are different for both.

Generally, a licensor who grants a nonexclusive license to copyrighted material waives the right to sue a licensee for copyright infringement, and may sue only for breach of contract. “If, however, a license is limited in scope and the licensee acts outside the scope, the licensor may bring an action for copyright infringement.” But how might a licensee act outside the scope of a license?

This chapter will present two cases, *Jacobsen v. Katzer* and *MDY v. Blizzard Entm’t*. The primary focus of these opinions is the difference between contractual covenants and contractual conditions. A covenant is an unqualified promise to perform or refrain from an act. A condition is an act or event, uncertain to occur, that must occur before a duty to perform arises. The Courts in these cases build upon the understanding that conditions define the scope of a contract, and use this for determining which acts of a licensee will therefore fall outside the scope of a contract and expose the licensee to the possibility of copyright infringement remedies.

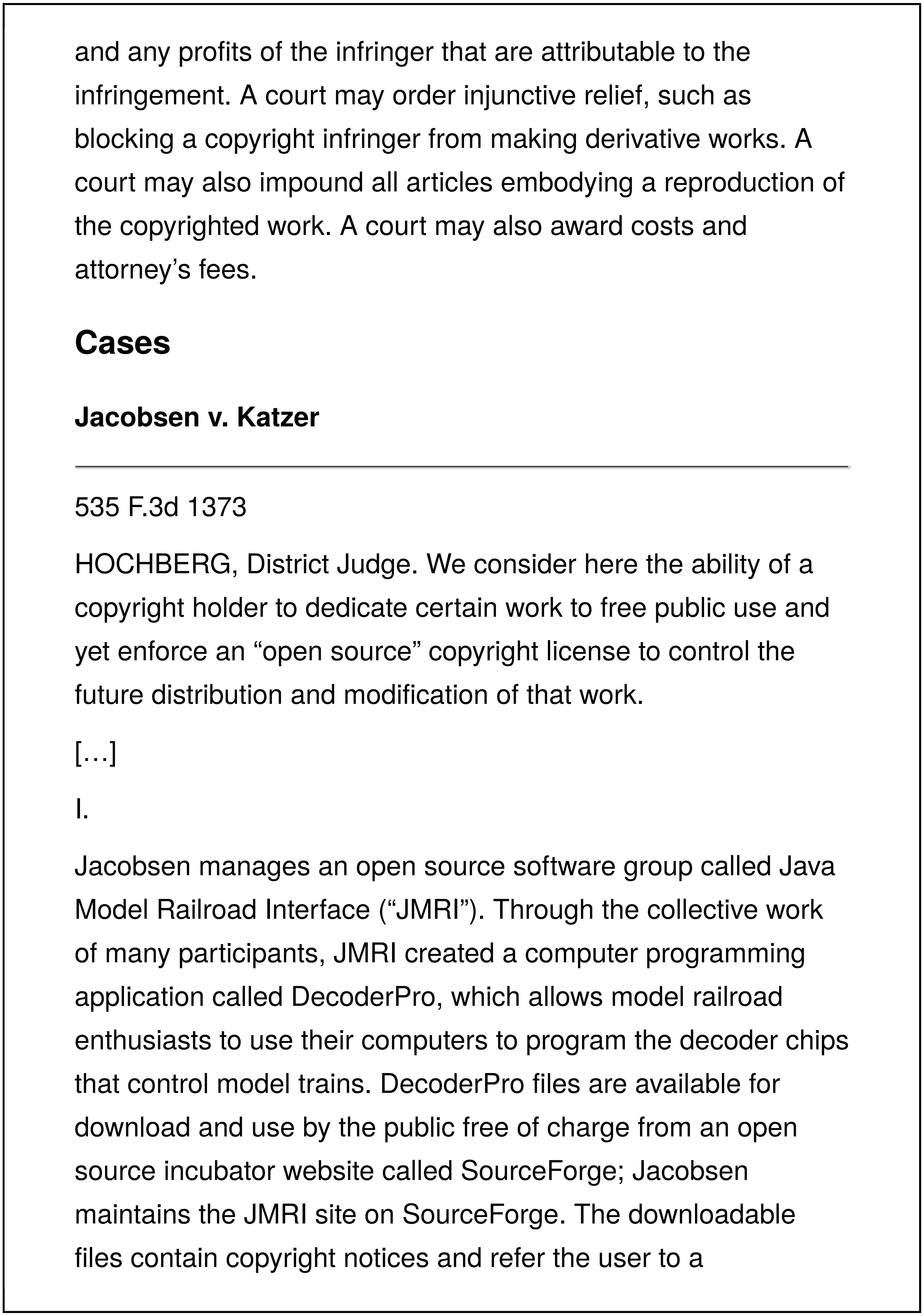
The *Jacobsen* court examines the connection between the copyright license at issue and the economic rights of the licensor. The *MDY* court considers whether the condition violated has a nexus to the licensor's exclusive rights of copyright. While reading these cases ask yourself, how are these tests related?

We commence with a comparison of the remedies for contract breach and copyright infringement to provide context. We conclude by reviewing the terms of a few open source licenses under the framework of covenants, conditions, and nexuses between conditions and the exclusive rights of copyright.

The Remedies Compared

Contract remedies are generally limited to an award of damages that will fulfill the "expectation interest" of the harmed party. In other words, the harmed party is owed the amount of money that will put them in as good a position as they would have been if the contract had been performed. Other available contract remedies include specific performance, for example injunctions to stop distributing a work of software, but specific performance will not be ordered if money is adequate to fulfill the harmed party's expectation interest. Punitive damages are unrecoverable for breach of contract unless the breach itself is a tort for which punitive damages are recoverable.

The remedies for copyright infringement, on the other hand, can include up to $150,000 in statutory damages per work infringed. Alternatively, a copyright holder can seek "to recover the actual damages suffered by him or her as a result of the infringement,



and any profits of the infringer that are attributable to the infringement. A court may order injunctive relief, such as blocking a copyright infringer from making derivative works. A court may also impound all articles embodying a reproduction of the copyrighted work. A court may also award costs and attorney’s fees.

Cases

Jacobsen v. Katzer

535 F.3d 1373

HOCHBERG, District Judge. We consider here the ability of a copyright holder to dedicate certain work to free public use and yet enforce an “open source” copyright license to control the future distribution and modification of that work.

[…]

I.

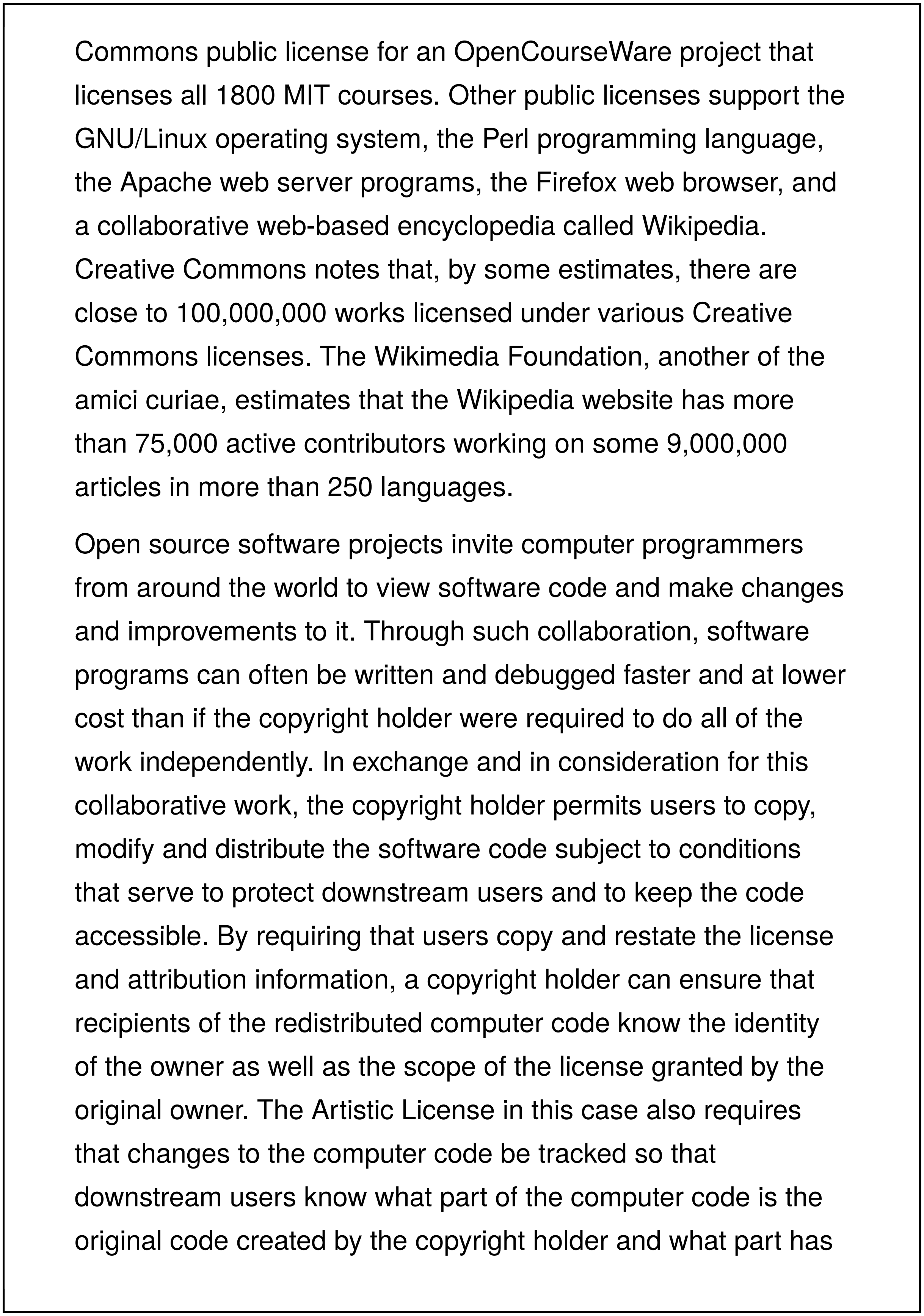
Jacobsen manages an open source software group called Java Model Railroad Interface (“JMRI”). Through the collective work of many participants, JMRI created a computer programming application called DecoderPro, which allows model railroad enthusiasts to use their computers to program the decoder chips that control model trains. DecoderPro files are available for download and use by the public free of charge from an open source incubator website called SourceForge; Jacobsen maintains the JMRI site on SourceForge. The downloadable files contain copyright notices and refer the user to a

“COPYING” file, which clearly sets forth the terms of the Artistic License.

Katzer/Kamind offers a competing software product, Decoder Commander, which is also used to program decoder chips. During development of Decoder Commander, one of Katzer/Kamind’s predecessors or employees is alleged to have downloaded the decoder definition files from DecoderPro and used portions of these files as part of the Decoder Commander software. The Decoder Commander software files that used DecoderPro definition files did not comply with the terms of the Artistic License. Specifically, the Decoder Commander software did not include (1) the authors’ names, (2) JMRI copyright notices, (3) references to the COPYING file, (4) an identification of SourceForge or JMRI as the original source of the definition files, and (5) a description of how the files or computer code had been changed from the original source code. The Decoder Commander software also changed various computer file names of DecoderPro files without providing a reference to the original JMRI files or information on where to get the Standard Version.

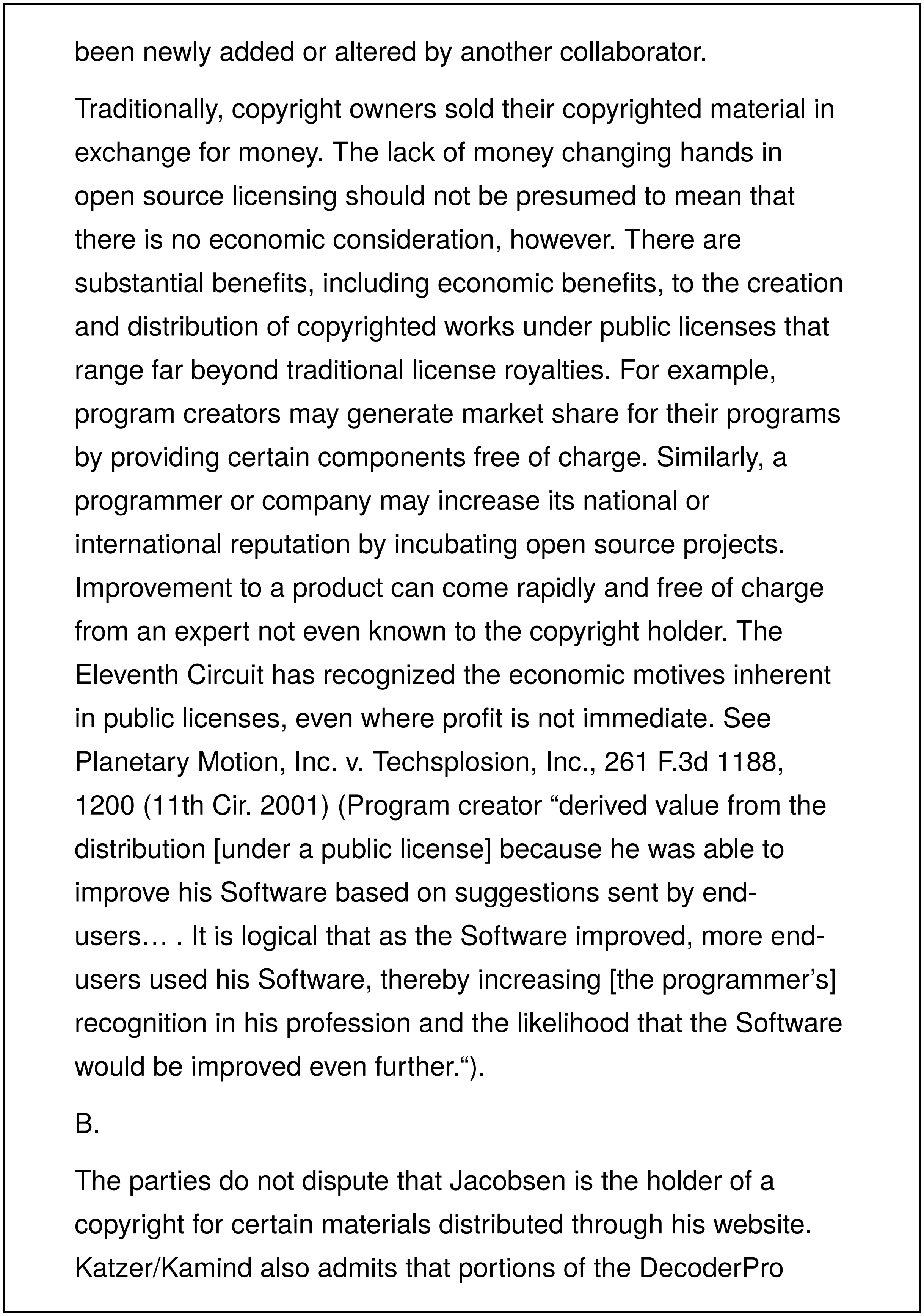
Jacobsen moved for a preliminary injunction, arguing that the violation of the terms of the Artistic License constituted copyright infringement and that, under Ninth Circuit law, irreparable harm could be presumed in a copyright infringement case. The District Court reviewed the Artistic License and determined that “Defendants’ alleged violation of the conditions of the license may have constituted a breach of the nonexclusive license, but does not create liability for copyright infringement where it would

> not otherwise exist.” [cite]. The District Court found that Jacobsen had a cause of action only for breach of contract, rather than an action for copyright infringement based on a breach of the conditions of the Artistic License. Because a breach of contract creates no presumption of irreparable harm, the District Court denied the motion for a preliminary injunction.
>
> Jacobsen appeals the finding that he does not have a cause of action for copyright infringement. Although an appeal concerning copyright law and not patent law is rare in our Circuit, here we indeed possess appellate jurisdiction.
>
> […]
>
> A.
>
> Public licenses, often referred to as “open source” licenses, are used by artists, authors, educators, software developers, and scientists who wish to create collaborative projects and to dedicate certain works to the public. Several types of public licenses have been designed to provide creators of copyrighted materials a means to protect and control their copyrights. Creative Commons, one of the amici curiae, provides free copyright licenses to allow parties to dedicate their works to the public or to license certain uses of their works while keeping some rights reserved.
>
> Open source licensing has become a widely used method of creative collaboration that serves to advance the arts and sciences in a manner and at a pace that few could have imagined just a few decades ago. For example, the Massachusetts Institute of Technology (“MIT”) uses a Creative



Commons public license for an OpenCourseWare project that licenses all 1800 MIT courses. Other public licenses support the GNU/Linux operating system, the Perl programming language, the Apache web server programs, the Firefox web browser, and a collaborative web-based encyclopedia called Wikipedia. Creative Commons notes that, by some estimates, there are close to 100,000,000 works licensed under various Creative Commons licenses. The Wikimedia Foundation, another of the amici curiae, estimates that the Wikipedia website has more than 75,000 active contributors working on some 9,000,000 articles in more than 250 languages.

Open source software projects invite computer programmers from around the world to view software code and make changes and improvements to it. Through such collaboration, software programs can often be written and debugged faster and at lower cost than if the copyright holder were required to do all of the work independently. In exchange and in consideration for this collaborative work, the copyright holder permits users to copy, modify and distribute the software code subject to conditions that serve to protect downstream users and to keep the code accessible. By requiring that users copy and restate the license and attribution information, a copyright holder can ensure that recipients of the redistributed computer code know the identity of the owner as well as the scope of the license granted by the original owner. The Artistic License in this case also requires that changes to the computer code be tracked so that downstream users know what part of the computer code is the original code created by the copyright holder and what part has



been newly added or altered by another collaborator.

Traditionally, copyright owners sold their copyrighted material in exchange for money. The lack of money changing hands in open source licensing should not be presumed to mean that there is no economic consideration, however. There are substantial benefits, including economic benefits, to the creation and distribution of copyrighted works under public licenses that range far beyond traditional license royalties. For example, program creators may generate market share for their programs by providing certain components free of charge. Similarly, a programmer or company may increase its national or international reputation by incubating open source projects. Improvement to a product can come rapidly and free of charge from an expert not even known to the copyright holder. The Eleventh Circuit has recognized the economic motives inherent in public licenses, even where profit is not immediate. See Planetary Motion, Inc. v. Techsplosion, Inc., 261 F.3d 1188, 1200 (11th Cir. 2001) (Program creator “derived value from the distribution [under a public license] because he was able to improve his Software based on suggestions sent by end-users… . It is logical that as the Software improved, more end-users used his Software, thereby increasing [the programmer’s] recognition in his profession and the likelihood that the Software would be improved even further.“).

B.

The parties do not dispute that Jacobsen is the holder of a copyright for certain materials distributed through his website. Katzer/Kamind also admits that portions of the DecoderPro software were copied, modified, and distributed as part of the Decoder Commander software. Accordingly, Jacobsen has made out a prima facie case of copyright infringement. Katzer/Kamind argues that they cannot be liable for copyright infringement because they had a license to use the material. Thus, the Court must evaluate whether the use by Katzer/Kamind was outside the scope of the license. *See LGS Architects*, 434 F.3d at 1156. The copyrighted materials in this case are downloadable by any user and are labeled to include a copyright notification and a COPYING file that includes the text of the Artistic License. The Artistic License grants users the right to copy, modify, and distribute the software:

provided that [the user] insert a prominent notice in each changed file stating how and when [the user] changed that file, and provided that [the user] do at least ONE of the following:

a) place [the user's] modifications in the Public Domain or otherwise make them Freely Available, such as by posting said modifications to Usenet or an equivalent medium, or placing the modifications on a major archive site such as ftp.uu.net, or by allowing the Copyright Holder to include [the user's] modifications in the Standard Version of the Package.

b) use the modified Package only within [the user's] corporation or organization.

c) rename any non-standard executables so the names do not conflict with the standard executables, which must also be provided, and provide a separate manual page for each nonstandard executable that clearly documents how it differs

> from the Standard Version, or
>
> d) make other distribution arrangements with the Copyright Holder.
>
> The heart of the argument on appeal concerns whether the terms of the Artistic License are conditions of, or merely covenants to, the copyright license. Generally, a "copyright owner who grants a nonexclusive license to use his copyrighted material waives his right to sue the licensee for copyright infringement" and can sue only for breach of contract. Sun Microsystems, Inc., v. Microsoft Corp., 188 F.3d 1115, 1121 (9th Cir. 1999); Graham v. James, 144 F.3d 229, 236 (2d Cir. 1998). If, however, a license is limited in scope and the licensee acts outside the scope, the licensor can bring an action for copyright infringement. See S.O.S., Inc. v. Payday, Inc., 886 F.2d 1081, 1087 (9th Cir.1989); Nimmer on Copyright, § 1015[A](1999).
>
> Thus, if the terms of the Artistic License allegedly violated are both covenants and conditions, they may serve to limit the scope of the license and are governed by copyright law. If they are merely covenants, by contrast, they are governed by contract law. *See Graham*, 144 F.3d at 236-37 whether breach of license is actionable as copyright infringement or breach of contract turns on whether provision breached is condition of the license, or mere covenant); Sun Microsystems, 188 F.3d at 1121 (following Graham; independent covenant does not limit scope of copyright license). The District Court did not expressly state whether the limitations in the Artistic License are independent covenants or, rather, conditions to the scope; its analysis, however, clearly treated the license limitations as contractual

covenants rather than conditions of the copyright license.

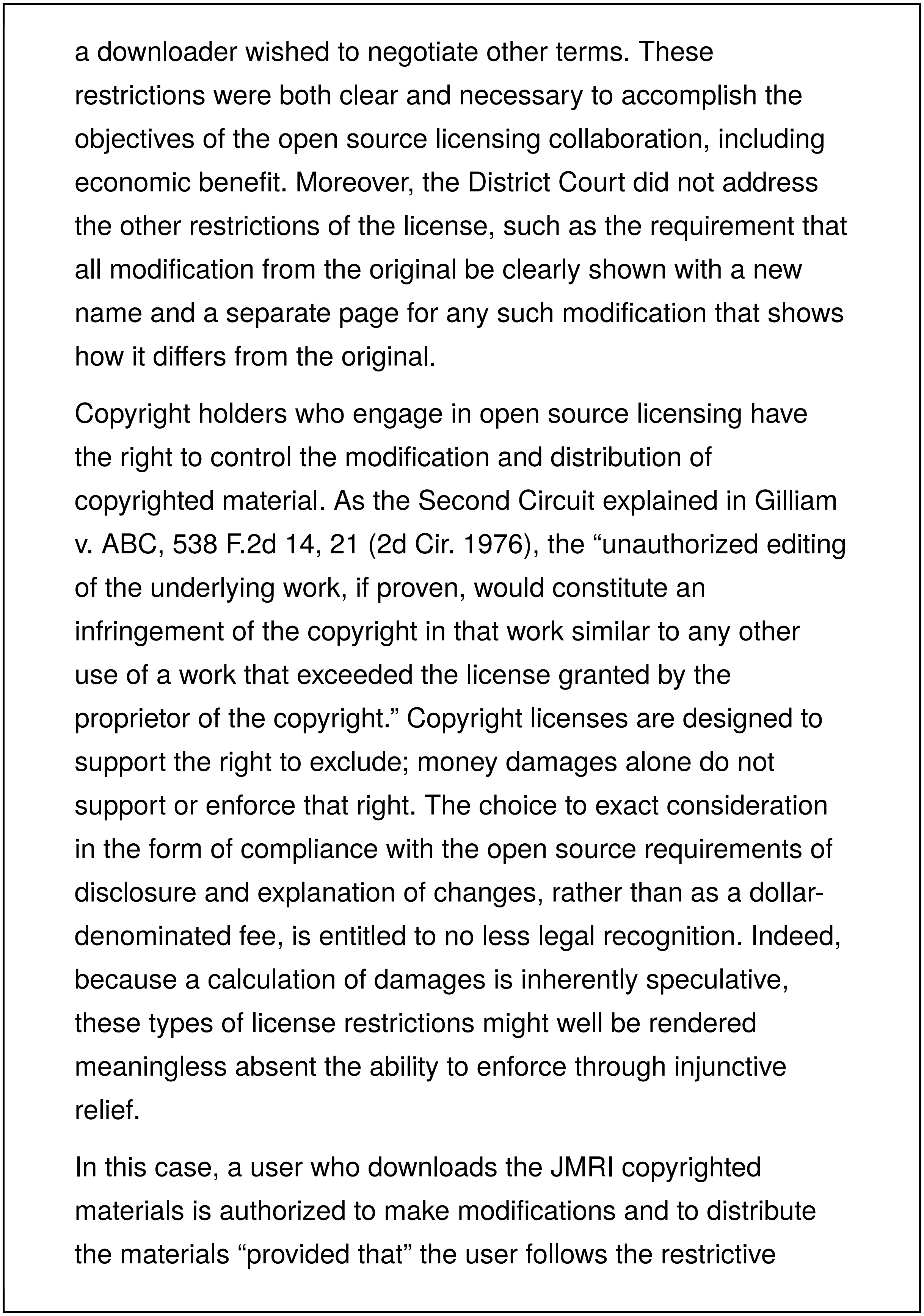
Jacobsen argues that the terms of the Artistic License define the scope of the license and that any use outside of these restrictions is copyright infringement. Katzer/Kamind argues that these terms do not limit the scope of the license and are merely covenants providing contractual terms for the use of the materials, and that his violation of them is neither compensable in damages nor subject to injunctive relief. Katzer/Kamind's argument is premised upon the assumption that Jacobsen's copyright gave him no economic rights because he made his computer code available to the public at no charge. From this assumption, Katzer/Kamind argues that copyright law does not recognize a cause of action for non-economic rights, relying on Gilliam v. ABC, 538 F.2d 14, 20-21 (2d Cir. 1976) "American copyright law, as presently written, does not recognize moral rights or provide a cause of action for their violation, since the law seeks to vindicate the economic, rather than the personal rights of authors.“). The District Court based its opinion on the breadth of the Artistic License terms, to which we now turn.

III.

The Artistic License states on its face that the document creates conditions: "The intent of this document is to state the conditions under which a Package may be copied." (Emphasis added.) The Artistic License also uses the traditional language of conditions by noting that the rights to copy, modify, and distribute are granted "provided that" the conditions are met. Under California contract law, "provided that" typically denotes a condition. *See, e.g., Diepenbrock v. Luiz*, 159 Cal. 716, 115 P. 743 (1911) (interpreting a real property lease reciting that when the property was sold, “this lease shall cease and be at an end, provided that the party of the first part shall then pay [certain compensation] to the party of the second part”; considering the appellant’s “interesting and ingenious” argument for interpreting this language as creating a mere covenant rather than a condition; and holding that this argument “cannot change the fact that, attributing the usual and ordinary signification to the language of the parties, a *condition* is found in the provision in question”) (emphases added).

The conditions set forth in the Artistic License are vital to enable the copyright holder to retain the ability to benefit from the work of downstream users. By requiring that users who modify or distribute the copyrighted material retain the reference to the original source files, downstream users are directed to Jacobsen’s website. Thus, downstream users know about the collaborative effort to improve and expand the SourceForge project once they learn of the “upstream” project from a “downstream” distribution, and they may join in that effort.

The District Court interpreted the Artistic License to permit a user to “modify the material in any way” and did not find that any of the “provided that” limitations in the Artistic License served to limit this grant. The District Court’s interpretation of the conditions of the Artistic License does not credit the explicit restrictions in the license that govern a downloader’s right to modify and distribute the copyrighted work. The copyright holder here expressly stated the terms upon which the right to modify and distribute the material depended and invited direct contact if



a downloader wished to negotiate other terms. These restrictions were both clear and necessary to accomplish the objectives of the open source licensing collaboration, including economic benefit. Moreover, the District Court did not address the other restrictions of the license, such as the requirement that all modification from the original be clearly shown with a new name and a separate page for any such modification that shows how it differs from the original.

Copyright holders who engage in open source licensing have the right to control the modification and distribution of copyrighted material. As the Second Circuit explained in Gilliam v. ABC, 538 F.2d 14, 21 (2d Cir. 1976), the "unauthorized editing of the underlying work, if proven, would constitute an infringement of the copyright in that work similar to any other use of a work that exceeded the license granted by the proprietor of the copyright." Copyright licenses are designed to support the right to exclude; money damages alone do not support or enforce that right. The choice to exact consideration in the form of compliance with the open source requirements of disclosure and explanation of changes, rather than as a dollar-denominated fee, is entitled to no less legal recognition. Indeed, because a calculation of damages is inherently speculative, these types of license restrictions might well be rendered meaningless absent the ability to enforce through injunctive relief.

In this case, a user who downloads the JMRI copyrighted materials is authorized to make modifications and to distribute the materials "provided that" the user follows the restrictive

> terms of the Artistic License. A copyright holder can grant the right to make certain modifications, yet retain his right to prevent other modifications. Indeed, such a goal is exactly the purpose of adding conditions to a license grant. The Artistic License, like many other common copyright licenses, requires that any copies that are distributed contain the copyright notices and the COPYING file. See, e.g., 3-10 Nimmer on Copyright § 10.15 (“An express (or possibly an implied) condition that a licensee must affix a proper copyright notice to all copies of the work that he causes to be published will render a publication devoid of such notice without authority from the licensor and therefore, an infringing act.“).
>
> It is outside the scope of the Artistic License to modify and distribute the copyrighted materials without copyright notices and a tracking of modifications from the original computer files. If a downloader does not assent to these conditions stated in the COPYING file, he is instructed to “make other arrangements with the Copyright Holder.” Katzer/Kamind did not make any such “other arrangements.” The clear language of the Artistic License creates conditions to protect the economic rights at issue in the granting of a public license. These conditions govern the rights to modify and distribute the computer programs and files included in the downloadable software package. The attribution and modification transparency requirements directly serve to drive traffic to the open source incubation page and to inform downstream users of the project, which is a significant economic goal of the copyright holder that the law will enforce. Through this controlled spread of

information, the copyright holder gains creative collaborators to the open source project; by requiring that changes made by downstream users be visible to the copyright holder and others, the copyright holder learns about the uses for his software and gains others' knowledge that can be used to advance future software releases.

IV.

For the aforementioned reasons, we vacate and remand. […] The judgment of the District Court is vacated and the case is remanded for further proceedings consistent with this opinion.

Discussion

1. The test for determining copyright liability for breach of a license stated by the Federal Circuit in *Jacobsen* relies on distinguishing covenants from conditions. This excerpt from a later case clarifies the distinction under California law:

```
"A covenant 'is another word for a contractual
promise.' A promise for
contract purposes 'is a manifestation of
intention to act or refrain from
acting in a specified way, so made as to
justify a promisee in understanding
that a commitment has been made.' Implied
covenants are disfavored and will
only be found if they effectuate the intent of
the parties, are a legal
necessity and 'after examining the contract as
```

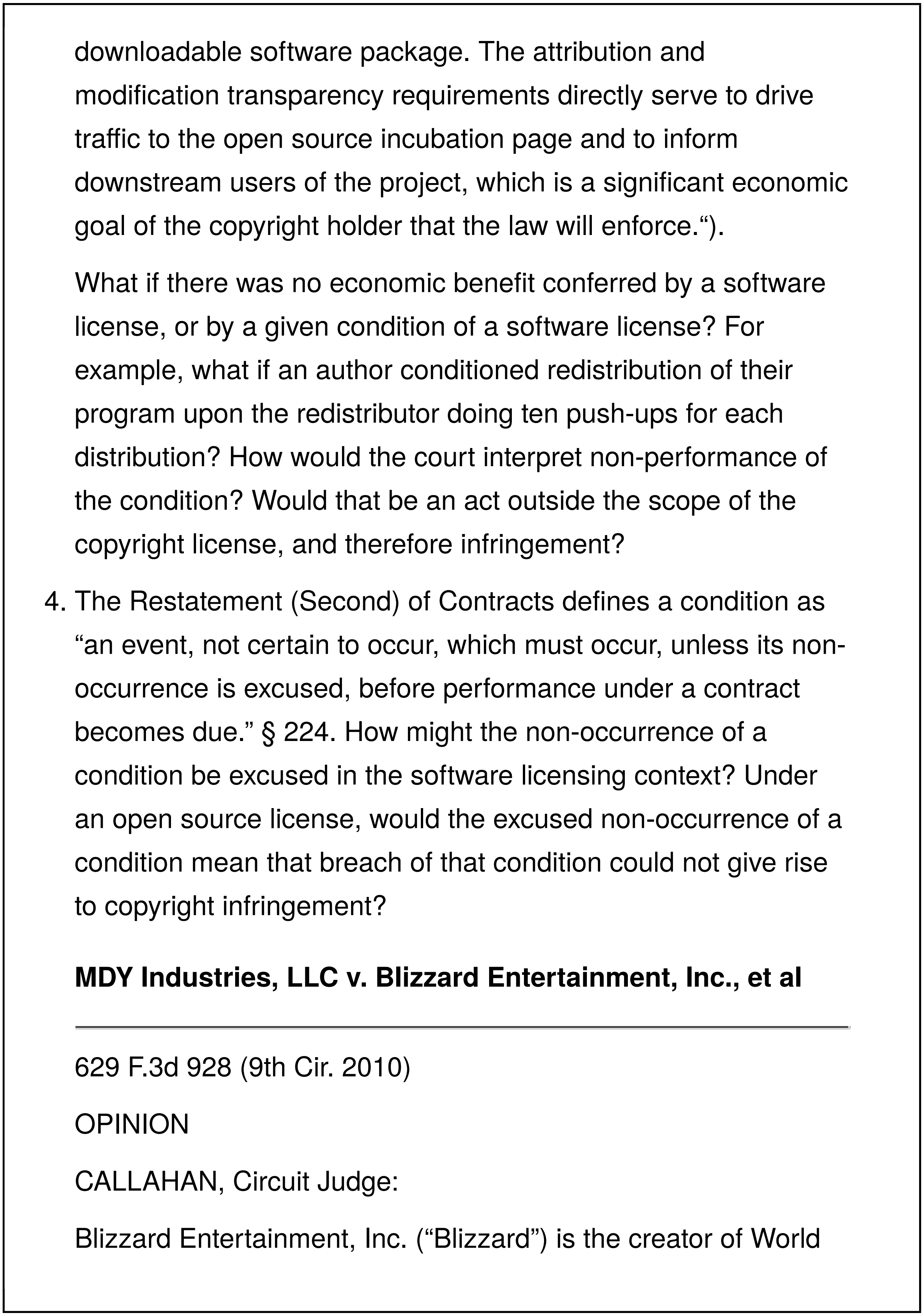
```
a whole it is [] obvious that
the parties had no reason to state the
covenant[.]' A condition, on the
other hand, 'is an event, not certain to occur,
which must occur, unless its
non-occurrence is excused, before performance
under a contract becomes due.'
Under California law a conditional obligation
is one 'when the rights or
duties of any party thereto depend upon the
occurrence of an uncertain
event.'" [Netbula, LLC v. Storage Tech. Corp.,
No. C06-07391 MJJ, 2008 U.S.
Dist. LEXIS 4119, at 9 (N.D. Cal. Jan. 17,
2008)](https://advance.lexis.com/api/document
/collection/cases/id/4RN5-M7P0-TXFP-
C34C-00000-00?page=9&reporter=1293&
context=1000516)*
(internal citations omitted).
```

2. When might a license term be a condition rather than a covenant?

   See Sun Microsystems, Inc. v. Microsoft Corp., 81 F. Supp. 2d 1026, 1032-33 (N.D. Cal. 2000)* (finding breached compatibility obligations to be a covenant, emphasizing that the license grant was not stated as being conditioned on the compatibility obligations and emphasizing the presence of a cure provision within the agreement); Montalvo v. LT's Benjamin Records, Inc., 56 F. Supp. 3d 121, 130 (D.P.R. 2014) (holding failure to pay royalties to be a breach of a covenant); *Sleash, LLC v. One Pet Planet, LLC, No. 3:14-cv-00863-ST, 2014 U.S. Dist. LEXIS 109253, at 51 (D. Or. Aug. 6, 2014) (finding one contract term to be a covenant where the term appeared in a separate section from the license grant and finding another contract term to be a covenant where the term was stated in promissory, as opposed to conditional, terms).

But see Accusoft Corp. v. Quest Diagnostics, Inc., No. 12-cv-40007-TSH, 2015 U.S. Dist. LEXIS 156693, at 76 (D. Mass. Aug. 19, 2015)* (finding a number of EULA terms to be conditions under Massachusetts law, as the terms clearly limited the scope of the license and were stated with emphatic, conditional language); Alaska Stock, LLC v. Pearson Educ., Inc., 975 F. Supp. 2d 1027, 1044 (D. Alaska 2013) (holding that a term limiting "the number of publications in which an image could appear [was] no different from the Ninth Circuit's example of the person who made a hundred copies of a book while licensed to make only one" and was therefore a condition); Jacobsen v. Katzer, 535 F.3d 1373, 1381 (Fed. Cir. 2008) ("Under California contract law, 'provided that' typically denotes a condition.").

3. The Jacobsen court recognizes that open source licensing schemes can confer an economic benefit upon licensors. Jacobsen v. Katzer, 535 F.3d 1373, 1382 (Fed. Cir. 2008)("The clear language of the Artistic License creates conditions to protect the economic rights at issue in the granting of a public license. These conditions govern the rights to modify and distribute the computer programs and files included in the



downloadable software package. The attribution and modification transparency requirements directly serve to drive traffic to the open source incubation page and to inform downstream users of the project, which is a significant economic goal of the copyright holder that the law will enforce.“).

What if there was no economic benefit conferred by a software license, or by a given condition of a software license? For example, what if an author conditioned redistribution of their program upon the redistributor doing ten push-ups for each distribution? How would the court interpret non-performance of the condition? Would that be an act outside the scope of the copyright license, and therefore infringement?

4. The Restatement (Second) of Contracts defines a condition as “an event, not certain to occur, which must occur, unless its non-occurrence is excused, before performance under a contract becomes due.” § 224. How might the non-occurrence of a condition be excused in the software licensing context? Under an open source license, would the excused non-occurrence of a condition mean that breach of that condition could not give rise to copyright infringement?

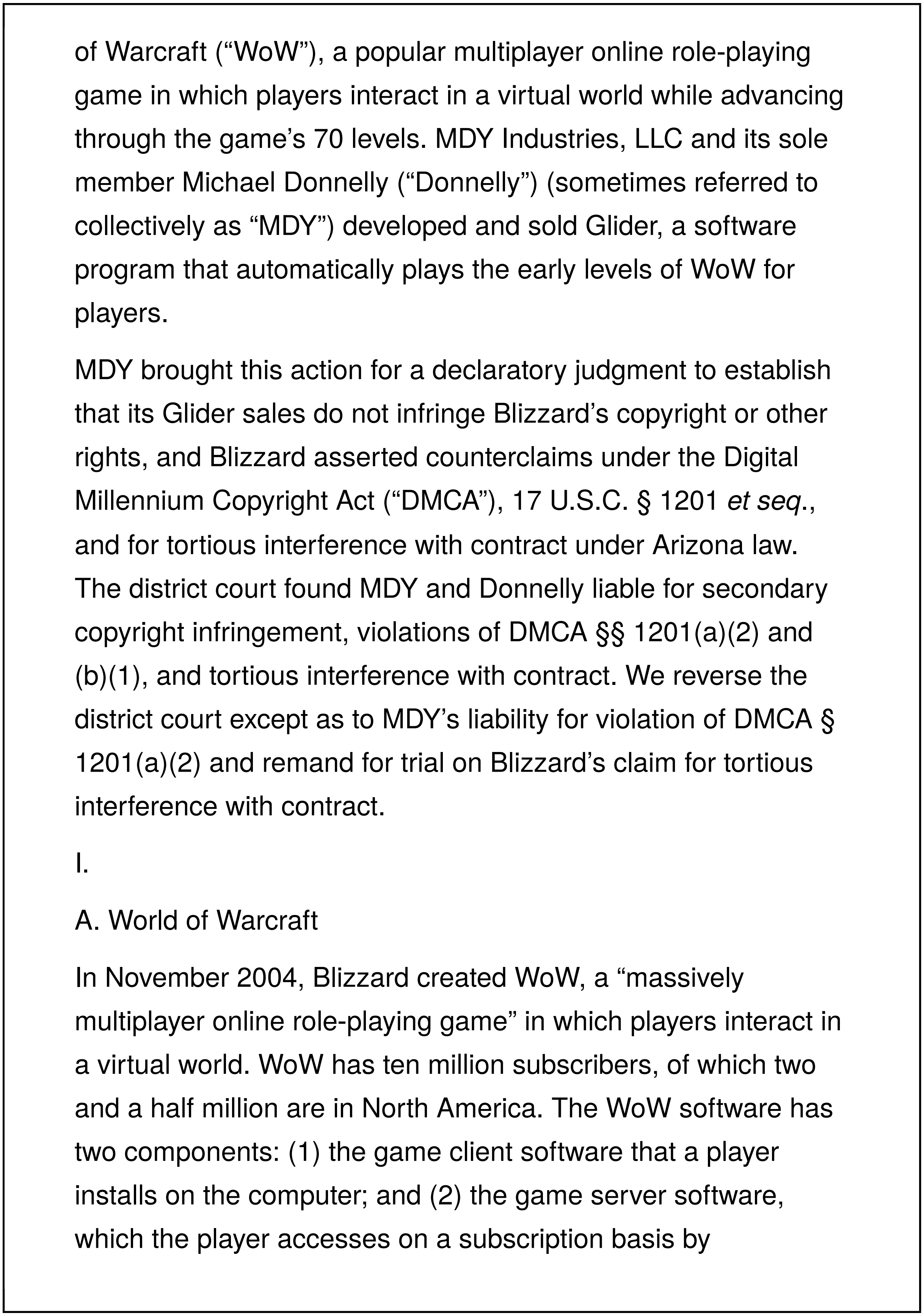
MDY Industries, LLC v. Blizzard Entertainment, Inc., et al

629 F.3d 928 (9th Cir. 2010)

OPINION

CALLAHAN, Circuit Judge:

Blizzard Entertainment, Inc. (“Blizzard”) is the creator of World



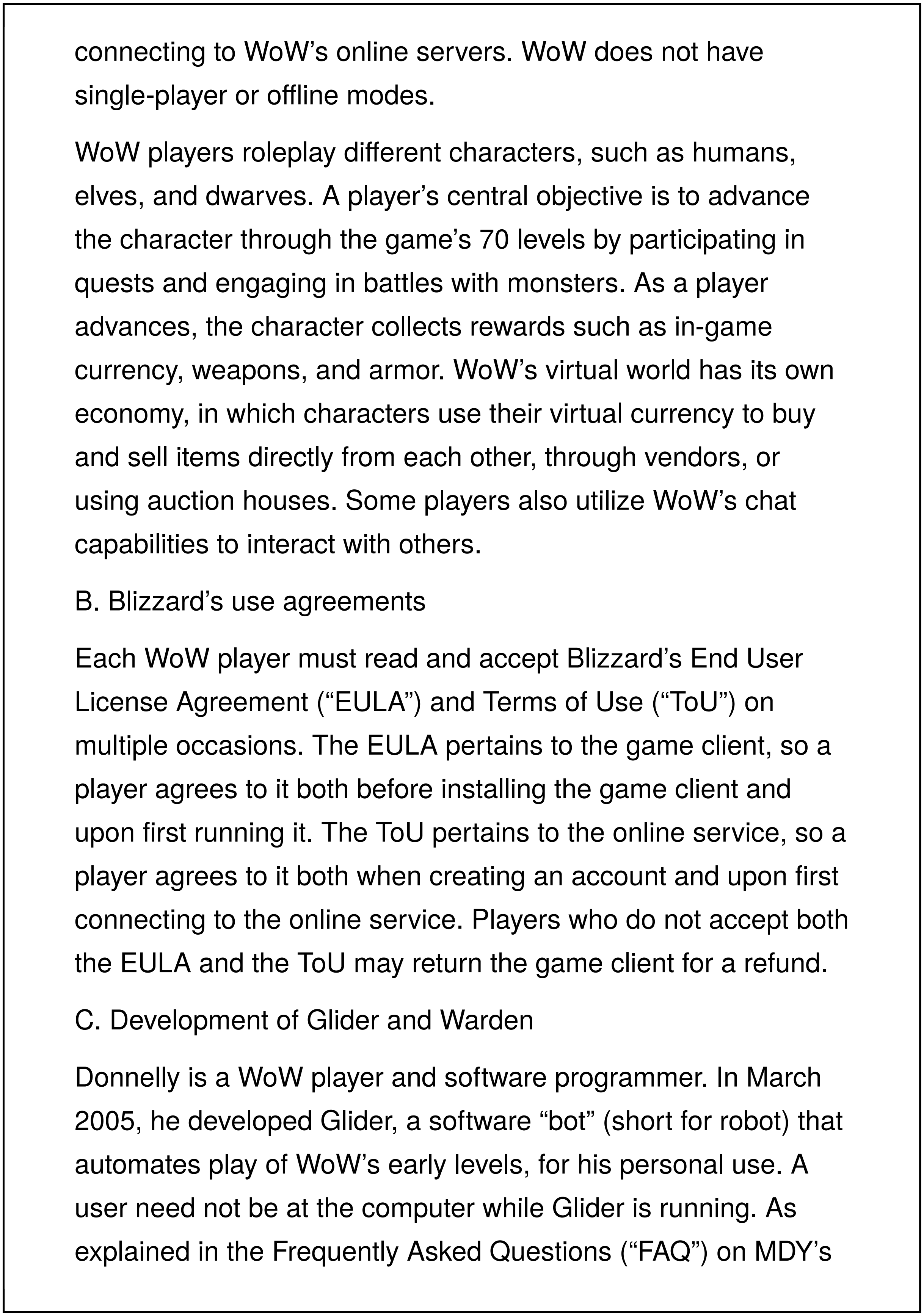
of Warcraft (“WoW”), a popular multiplayer online role-playing game in which players interact in a virtual world while advancing through the game’s 70 levels. MDY Industries, LLC and its sole member Michael Donnelly (“Donnelly”) (sometimes referred to collectively as “MDY”) developed and sold Glider, a software program that automatically plays the early levels of WoW for players.

MDY brought this action for a declaratory judgment to establish that its Glider sales do not infringe Blizzard’s copyright or other rights, and Blizzard asserted counterclaims under the Digital Millennium Copyright Act (“DMCA”), 17 U.S.C. § 1201 *et seq*., and for tortious interference with contract under Arizona law. The district court found MDY and Donnelly liable for secondary copyright infringement, violations of DMCA §§ 1201(a)(2) and (b)(1), and tortious interference with contract. We reverse the district court except as to MDY’s liability for violation of DMCA § 1201(a)(2) and remand for trial on Blizzard’s claim for tortious interference with contract.

I.

A. World of Warcraft

In November 2004, Blizzard created WoW, a “massively multiplayer online role-playing game” in which players interact in a virtual world. WoW has ten million subscribers, of which two and a half million are in North America. The WoW software has two components: (1) the game client software that a player installs on the computer; and (2) the game server software, which the player accesses on a subscription basis by



connecting to WoW's online servers. WoW does not have single-player or offline modes.

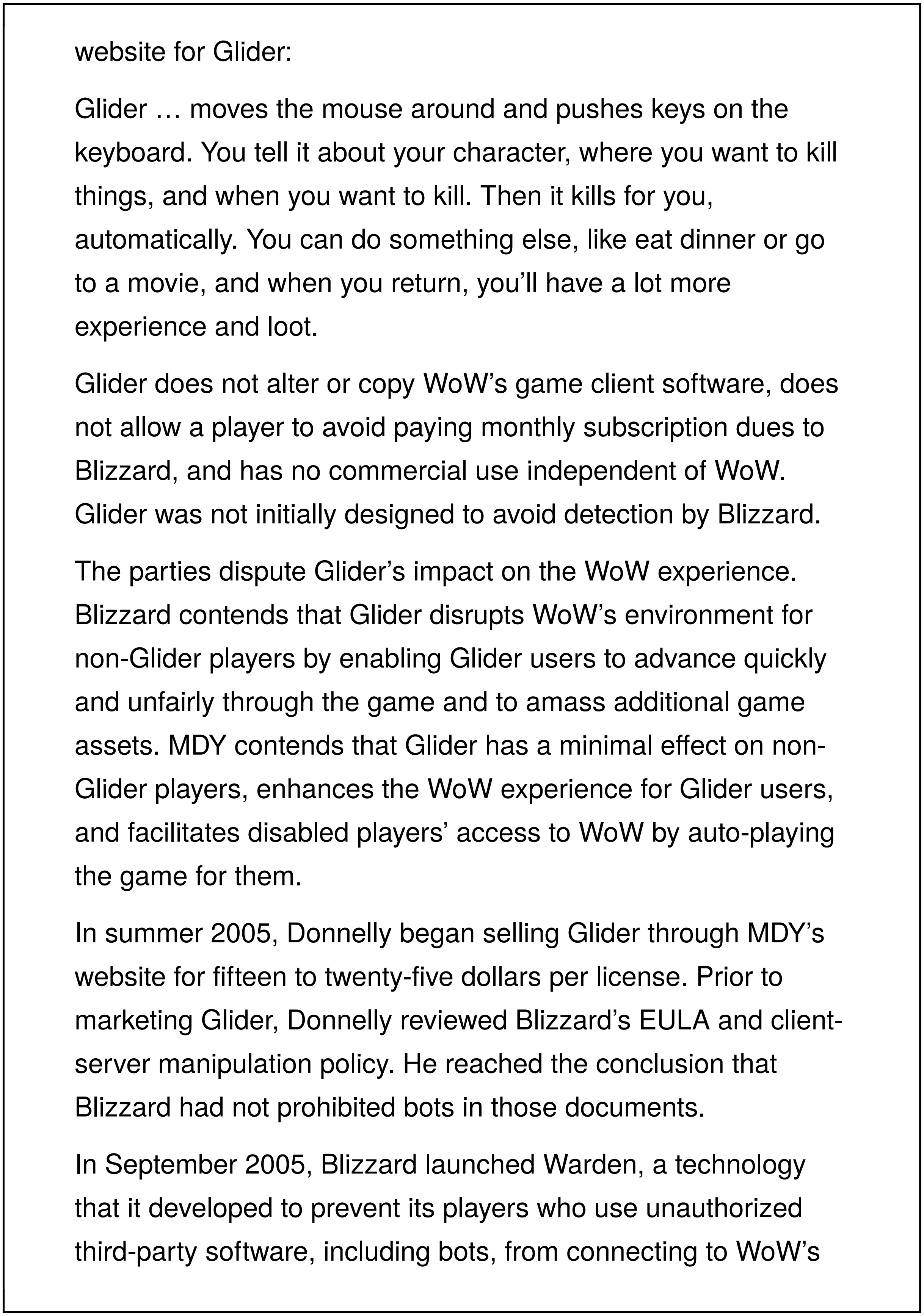
WoW players roleplay different characters, such as humans, elves, and dwarves. A player's central objective is to advance the character through the game's 70 levels by participating in quests and engaging in battles with monsters. As a player advances, the character collects rewards such as in-game currency, weapons, and armor. WoW's virtual world has its own economy, in which characters use their virtual currency to buy and sell items directly from each other, through vendors, or using auction houses. Some players also utilize WoW's chat capabilities to interact with others.

B. Blizzard's use agreements

Each WoW player must read and accept Blizzard's End User License Agreement ("EULA") and Terms of Use ("ToU") on multiple occasions. The EULA pertains to the game client, so a player agrees to it both before installing the game client and upon first running it. The ToU pertains to the online service, so a player agrees to it both when creating an account and upon first connecting to the online service. Players who do not accept both the EULA and the ToU may return the game client for a refund.

C. Development of Glider and Warden

Donnelly is a WoW player and software programmer. In March 2005, he developed Glider, a software "bot" (short for robot) that automates play of WoW's early levels, for his personal use. A user need not be at the computer while Glider is running. As explained in the Frequently Asked Questions ("FAQ") on MDY's



website for Glider:

Glider … moves the mouse around and pushes keys on the keyboard. You tell it about your character, where you want to kill things, and when you want to kill. Then it kills for you, automatically. You can do something else, like eat dinner or go to a movie, and when you return, you'll have a lot more experience and loot.

Glider does not alter or copy WoW's game client software, does not allow a player to avoid paying monthly subscription dues to Blizzard, and has no commercial use independent of WoW. Glider was not initially designed to avoid detection by Blizzard.

The parties dispute Glider's impact on the WoW experience. Blizzard contends that Glider disrupts WoW's environment for non-Glider players by enabling Glider users to advance quickly and unfairly through the game and to amass additional game assets. MDY contends that Glider has a minimal effect on non-Glider players, enhances the WoW experience for Glider users, and facilitates disabled players' access to WoW by auto-playing the game for them.

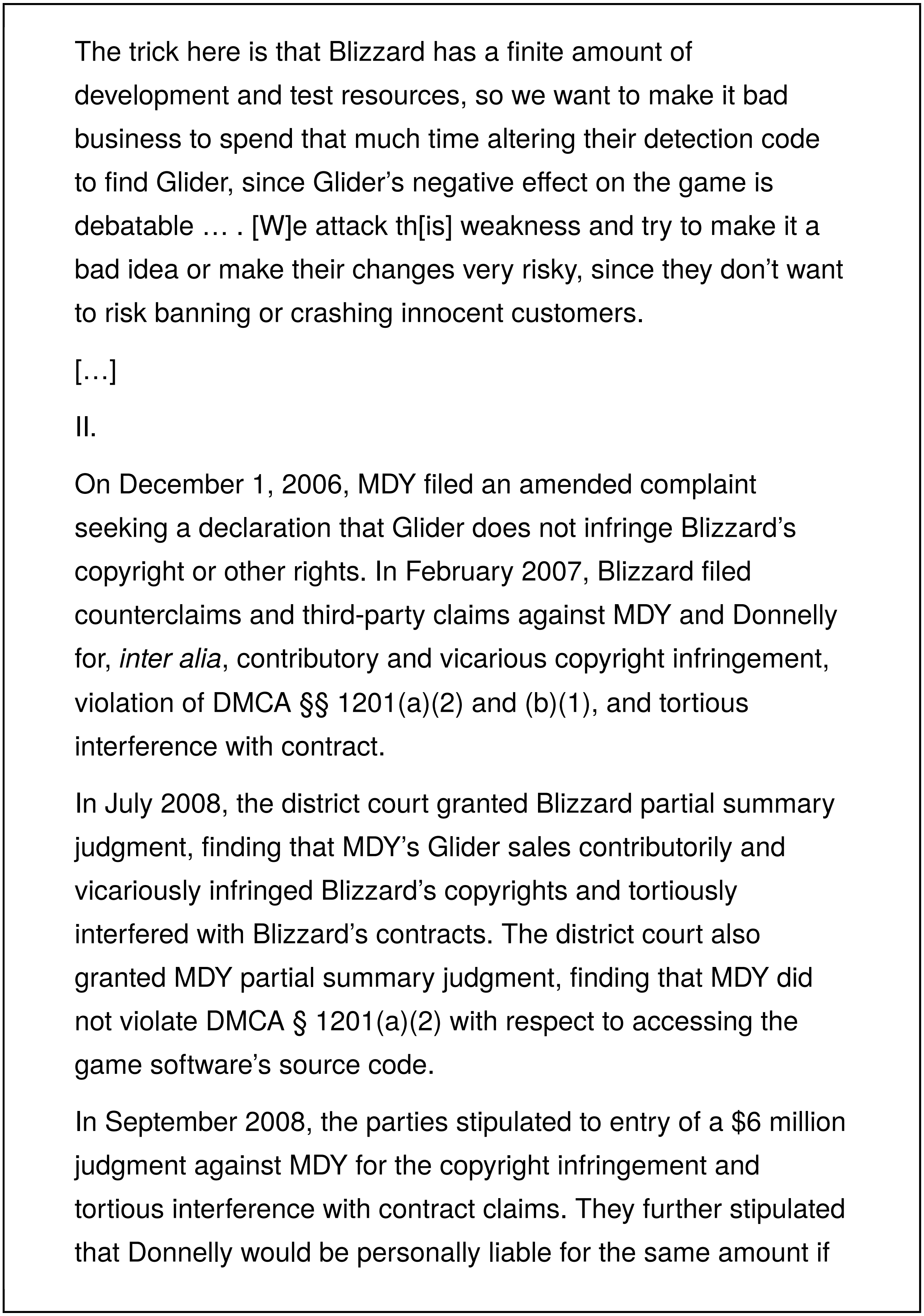
In summer 2005, Donnelly began selling Glider through MDY's website for fifteen to twenty-five dollars per license. Prior to marketing Glider, Donnelly reviewed Blizzard's EULA and client-server manipulation policy. He reached the conclusion that Blizzard had not prohibited bots in those documents.

In September 2005, Blizzard launched Warden, a technology that it developed to prevent its players who use unauthorized third-party software, including bots, from connecting to WoW's servers. Warden was able to detect Glider, and Blizzard immediately used Warden to ban most Glider users. MDY responded by modifying Glider to avoid detection and promoting its new anti-detection features on its website's FAQ. It added a subscription service, Glider Elite, which offered "additional protection from game detection software" for five dollars a month.

Thus, by late 2005, MDY was aware that Blizzard was prohibiting bots. MDY modified its website to indicate that using Glider violated Blizzard's ToU. In November 2005, Donnelly wrote in an email interview, "Avoiding detection is rather exciting, to be sure. Since Blizzard does not want bots running at all, it's a violation to use them." Following MDY's anti-detection modifications, Warden only occasionally detected Glider. As of September 2008, MDY had gross revenues of $3.5 million based on 120,000 Glider license sales.

D. Financial and practical impact of Glider

Blizzard claims that from December 2004 to March 2008, it received 465,000 complaints about WoW bots, several thousand of which named Glider. Blizzard spends $940,000 annually to respond to these complaints, and the parties have stipulated that Glider is the principal bot used by WoW players. Blizzard introduced evidence that it may have lost monthly subscription fees from Glider users, who were able to reach WoW's highest levels in fewer weeks than players playing manually. Donnelly acknowledged in a November 2005 email that MDY's business strategy was to make Blizzard's anti-bot detection attempts financially prohibitive:



The trick here is that Blizzard has a finite amount of development and test resources, so we want to make it bad business to spend that much time altering their detection code to find Glider, since Glider's negative effect on the game is debatable … . [W]e attack th[is] weakness and try to make it a bad idea or make their changes very risky, since they don't want to risk banning or crashing innocent customers.

[…]

II.

On December 1, 2006, MDY filed an amended complaint seeking a declaration that Glider does not infringe Blizzard's copyright or other rights. In February 2007, Blizzard filed counterclaims and third-party claims against MDY and Donnelly for, *inter alia*, contributory and vicarious copyright infringement, violation of DMCA §§ 1201(a)(2) and (b)(1), and tortious interference with contract.

In July 2008, the district court granted Blizzard partial summary judgment, finding that MDY's Glider sales contributorily and vicariously infringed Blizzard's copyrights and tortiously interfered with Blizzard's contracts. The district court also granted MDY partial summary judgment, finding that MDY did not violate DMCA § 1201(a)(2) with respect to accessing the game software's source code.

In September 2008, the parties stipulated to entry of a $6 million judgment against MDY for the copyright infringement and tortious interference with contract claims. They further stipulated that Donnelly would be personally liable for the same amount if found personally liable at trial. After a January 2009 bench trial, the district court held MDY liable under DMCA §§ 1201(a)(2) and (b)(1). It also held Donnelly personally liable for MDY's copyright infringement, DMCA violations, and tortious interference with contract.

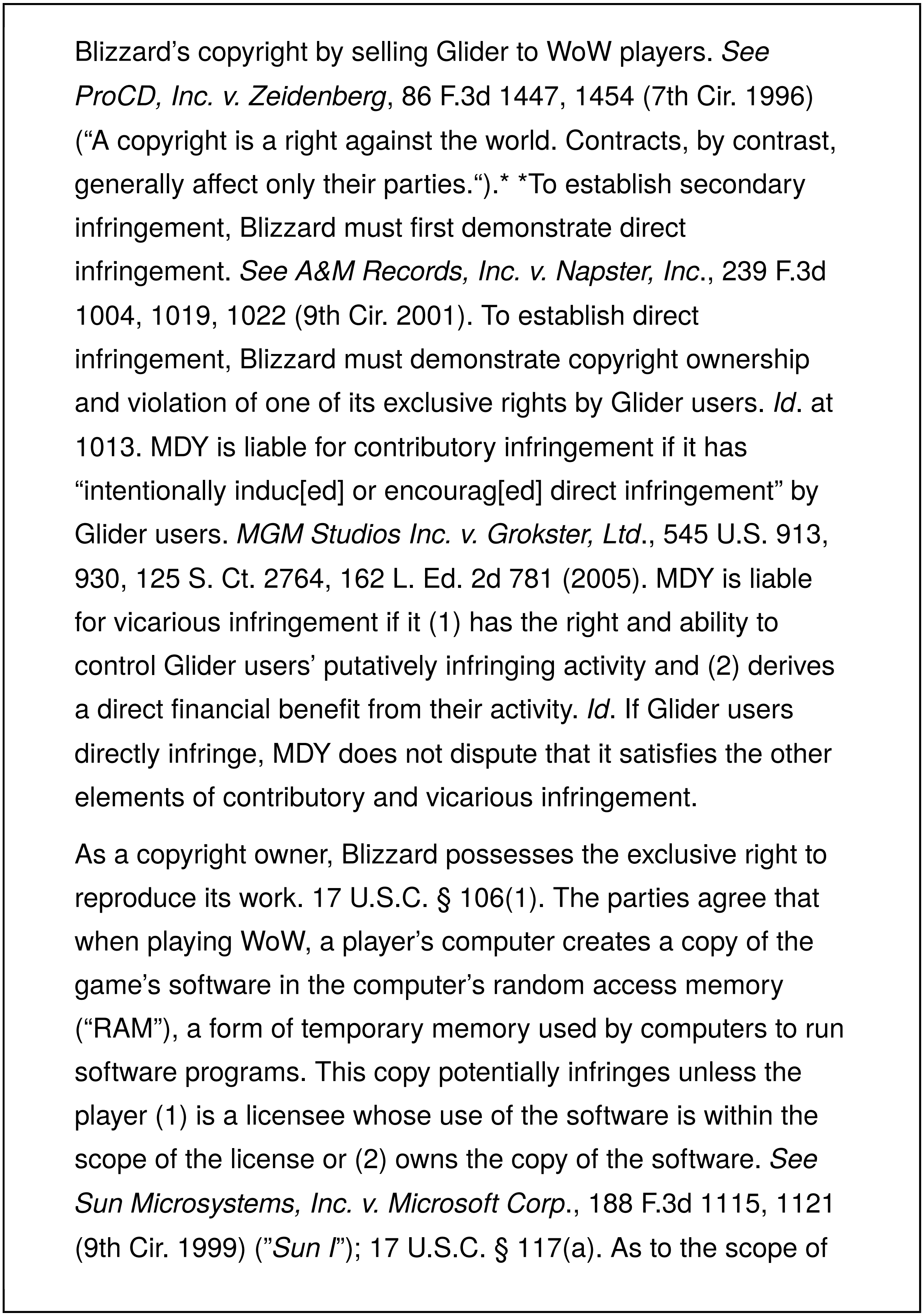
On April 1, 2009, the district court entered judgment against MDY and Donnelly for $6.5 million, an adjusted figure to which the parties stipulated based on MDY's DMCA liability and post-summary judgment Glider sales. The district court permanently enjoined MDY from distributing Glider. MDY's efforts to stay injunctive relief pending appeal were unsuccessful. On April 29, 2009, MDY timely filed this appeal. On May 12, 2009, Blizzard timely cross-appealed the district court's holding that MDY did not violate DMCA §§ 1201(a)(2) and (b)(1) as to the game software's source code.

III.

We review de novo the district court's (1) orders granting or denying summary judgment; (2) conclusions of law after a bench trial; and (3) interpretations of state law. *Padfield v. AIG Life Ins*., 290 F.3d 1121, 1124 (9th Cir. 2002); *Twentieth Century Fox Film Corp. v. Entm't Distrib*., 429 F.3d 869, 879 (9th Cir. 2005); *Laws v. Sony Music Entm't, Inc*., 448 F.3d 1134, 1137 (9th Cir. 2006). We review the district court's findings of fact for clear error. *Twentieth Century Fox*, 429 F.3d at 879.

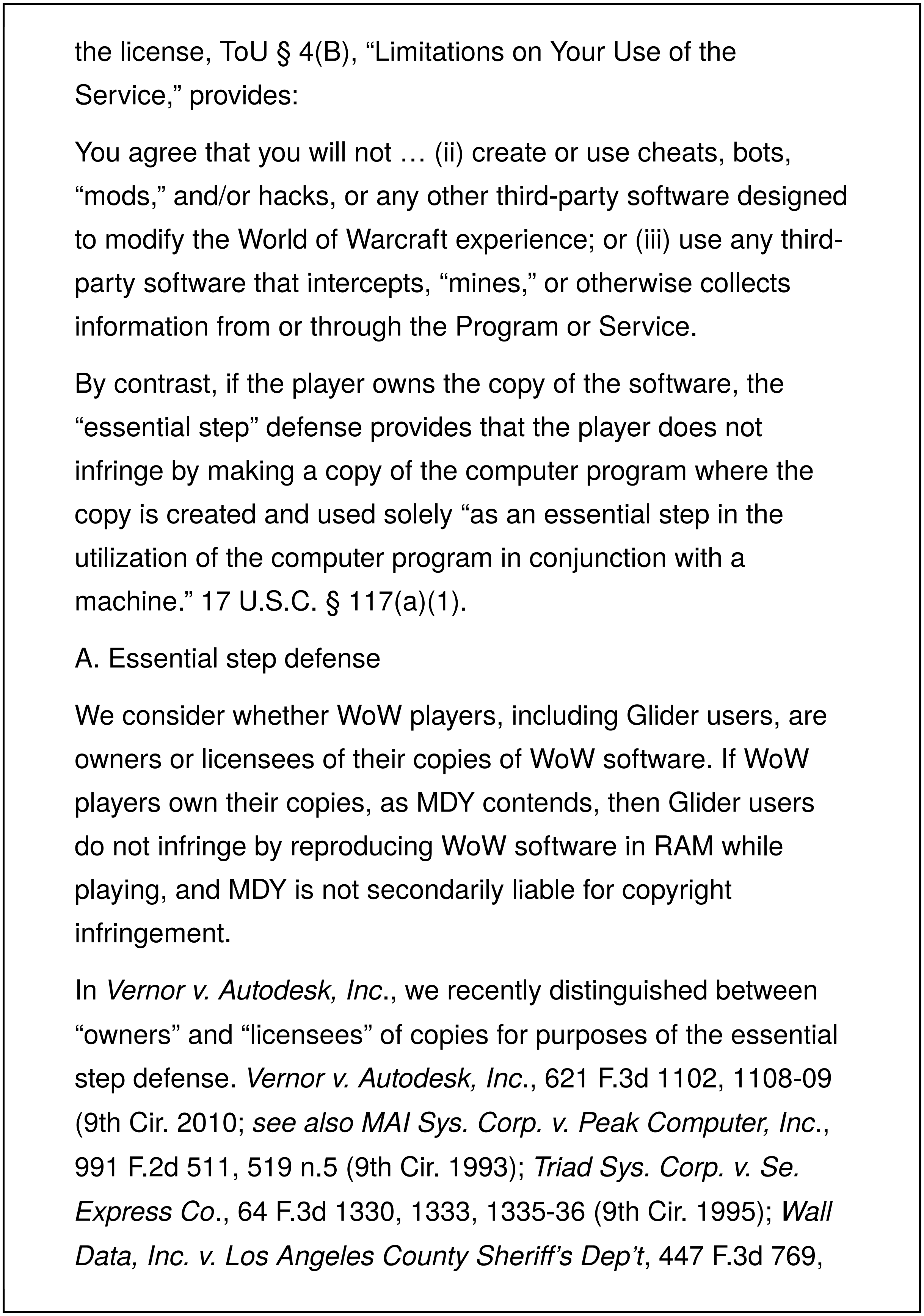
IV.

We first consider whether MDY committed contributory or vicarious infringement (collectively, "secondary infringement") of



Blizzard's copyright by selling Glider to WoW players. *See ProCD, Inc. v. Zeidenberg*, 86 F.3d 1447, 1454 (7th Cir. 1996) ("A copyright is a right against the world. Contracts, by contrast, generally affect only their parties.").* *To establish secondary infringement, Blizzard must first demonstrate direct infringement. *See A&M Records, Inc. v. Napster, Inc.*, 239 F.3d 1004, 1019, 1022 (9th Cir. 2001). To establish direct infringement, Blizzard must demonstrate copyright ownership and violation of one of its exclusive rights by Glider users. *Id*. at 1013. MDY is liable for contributory infringement if it has "intentionally induc[ed] or encourag[ed] direct infringement" by Glider users. *MGM Studios Inc. v. Grokster, Ltd.*, 545 U.S. 913, 930, 125 S. Ct. 2764, 162 L. Ed. 2d 781 (2005). MDY is liable for vicarious infringement if it (1) has the right and ability to control Glider users' putatively infringing activity and (2) derives a direct financial benefit from their activity. *Id*. If Glider users directly infringe, MDY does not dispute that it satisfies the other elements of contributory and vicarious infringement.

As a copyright owner, Blizzard possesses the exclusive right to reproduce its work. 17 U.S.C. § 106(1). The parties agree that when playing WoW, a player's computer creates a copy of the game's software in the computer's random access memory ("RAM"), a form of temporary memory used by computers to run software programs. This copy potentially infringes unless the player (1) is a licensee whose use of the software is within the scope of the license or (2) owns the copy of the software. *See Sun Microsystems, Inc. v. Microsoft Corp.*, 188 F.3d 1115, 1121 (9th Cir. 1999) ("*Sun I*"); 17 U.S.C. § 117(a). As to the scope of



the license, ToU § 4(B), “Limitations on Your Use of the Service,” provides:

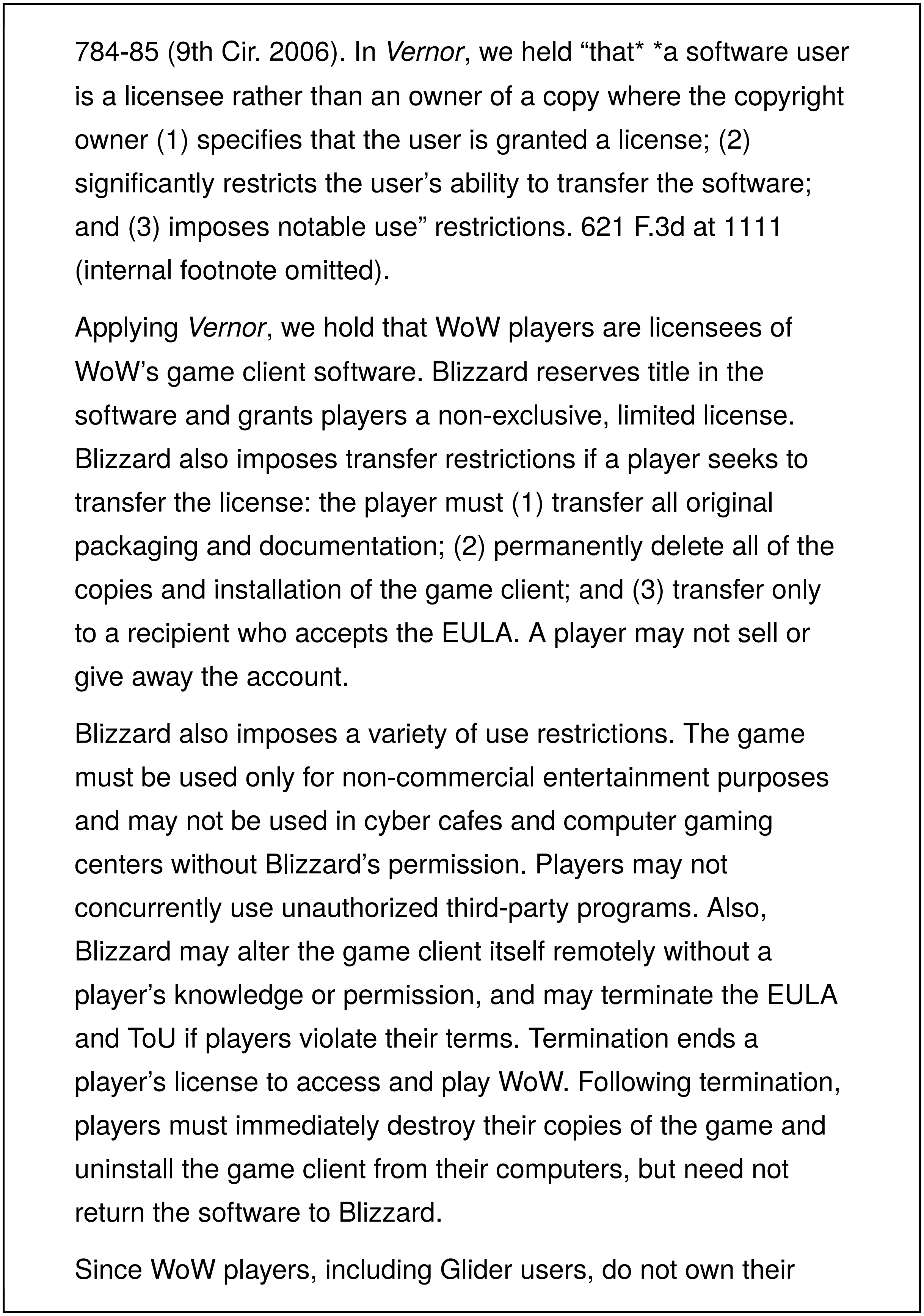
You agree that you will not … (ii) create or use cheats, bots, “mods,” and/or hacks, or any other third-party software designed to modify the World of Warcraft experience; or (iii) use any third-party software that intercepts, “mines,” or otherwise collects information from or through the Program or Service.

By contrast, if the player owns the copy of the software, the “essential step” defense provides that the player does not infringe by making a copy of the computer program where the copy is created and used solely “as an essential step in the utilization of the computer program in conjunction with a machine.” 17 U.S.C. § 117(a)(1).

A. Essential step defense

We consider whether WoW players, including Glider users, are owners or licensees of their copies of WoW software. If WoW players own their copies, as MDY contends, then Glider users do not infringe by reproducing WoW software in RAM while playing, and MDY is not secondarily liable for copyright infringement.

In *Vernor v. Autodesk, Inc.*, we recently distinguished between “owners” and “licensees” of copies for purposes of the essential step defense. *Vernor v. Autodesk, Inc.*, 621 F.3d 1102, 1108-09 (9th Cir. 2010; *see also MAI Sys. Corp. v. Peak Computer, Inc.*, 991 F.2d 511, 519 n.5 (9th Cir. 1993); *Triad Sys. Corp. v. Se. Express Co.*, 64 F.3d 1330, 1333, 1335-36 (9th Cir. 1995); *Wall Data, Inc. v. Los Angeles County Sheriff’s Dep’t*, 447 F.3d 769,



784-85 (9th Cir. 2006). In *Vernor*, we held “that* *a software user is a licensee rather than an owner of a copy where the copyright owner (1) specifies that the user is granted a license; (2) significantly restricts the user’s ability to transfer the software; and (3) imposes notable use” restrictions. 621 F.3d at 1111 (internal footnote omitted).

Applying *Vernor*, we hold that WoW players are licensees of WoW’s game client software. Blizzard reserves title in the software and grants players a non-exclusive, limited license. Blizzard also imposes transfer restrictions if a player seeks to transfer the license: the player must (1) transfer all original packaging and documentation; (2) permanently delete all of the copies and installation of the game client; and (3) transfer only to a recipient who accepts the EULA. A player may not sell or give away the account.

Blizzard also imposes a variety of use restrictions. The game must be used only for non-commercial entertainment purposes and may not be used in cyber cafes and computer gaming centers without Blizzard’s permission. Players may not concurrently use unauthorized third-party programs. Also, Blizzard may alter the game client itself remotely without a player’s knowledge or permission, and may terminate the EULA and ToU if players violate their terms. Termination ends a player’s license to access and play WoW. Following termination, players must immediately destroy their copies of the game and uninstall the game client from their computers, but need not return the software to Blizzard.

Since WoW players, including Glider users, do not own their copies of the software, Glider users may not claim the essential step defense. 17 U.S.C. § 117(a)(1). Thus, when their computers copy WoW software into RAM, the players may infringe unless their usage is within the scope of Blizzard's limited license.

B. Contractual covenants vs. license conditions

"A copyright owner who grants a nonexclusive, limited license ordinarily waives the right to sue licensees for copyright infringement, and it may sue only for breach of contract." *Sun I*, 188 F.3d at 1121 (internal quotations omitted). However, if the licensee acts outside the scope of the license, the licensor may sue for copyright infringement. *Id*. (citing *S.O.S., Inc. v. Payday, Inc*., 886 F.2d 1081, 1087 (9th Cir. 1989)). Enforcing a copyright license "raises issues that lie at the intersection of copyright and contract law." *Id*. at 1122.

We refer to contractual terms that limit a license's scope as "conditions," the breach of which constitute copyright infringement. *Id*. at 1120. We refer to all other license terms as "covenants," the breach of which is actionable only under contract law. *Id*. We distinguish between conditions and covenants according to state contract law, to the extent consistent with federal copyright law and policy. *Foad Consulting Group v. Musil Govan Azzalino*, 270 F.3d 821, 827 (9th Cir. 2001).

A Glider user commits copyright infringement by playing WoW while violating a ToU term that is a license condition. To establish copyright infringement, then, Blizzard must demonstrate that the violated term — ToU § 4(B) — is a condition rather than a covenant. *Sun I*, 188 F.3d at 1122. Blizzard's EULAs and ToUs provide that they are to be interpreted according to Delaware law. Accordingly, we first construe them under Delaware law, and then evaluate whether that construction is consistent with federal copyright law and policy.

A covenant is a contractual promise, i.e., a manifestation of intention to act or refrain from acting in a particular way, such that the promisee is justified in understanding that the promisor has made a commitment. *See Travel Centers of Am. LLC v. Brog*, No. 3751-CC, 2008 Del. Ch. LEXIS 183, *9 (Del. Ch. Dec. 5, 2008); *see also* Restatement (Second) of Contracts § 2 (1981). A condition precedent is an act or event that must occur before a duty to perform a promise arises. *AES P.R., L.P. v. Alstom Power, Inc.*, 429 F. Supp. 2d 713, 717 (D. Del. 2006) (citing Delaware state law); *see also* Restatement (Second) of Contracts § 224. Conditions precedent are disfavored because they tend to work forfeitures. *AES*, 429 F. Supp. 2d at 717 (internal citations omitted). Wherever possible, equity construes ambiguous contract provisions as covenants rather than conditions. See* Wilmington Tr. Co. v. Clark*, 325 A.2d 383, 386 (Del. Ch. 1974). However, if the contract is unambiguous, the court construes it according to its terms. *AES*, 429 F. Supp. 2d at 717 (citing 17 Am. Jur. 2d Contracts § 460 (2006)).

Applying these principles, ToU § 4(B)(ii) and (iii)'s prohibitions against bots and unauthorized third-party software are covenants rather than copyright-enforceable conditions. *See*

Greenwood v. CompuCredit Corp., 615 F.3d 1204, 1212, (9th Cir. 2010) ("[H]eadings and titles are not meant to take the place of the detailed provisions of the text," and … "the heading of a section cannot limit the plain meaning of the text." (quoting *Bhd. of R.R. Trainmen v. Balt. & Ohio R.R*., 331 U.S. 519, 528-29, 67 S. Ct. 1387, 91 L. Ed. 1646 (1947))). Although ToU § 4 is titled, "Limitations on Your Use of the Service," nothing in that section conditions Blizzard's grant of a limited license on players' compliance with ToU § 4's restrictions. To the extent that the title introduces any ambiguity, under Delaware law, ToU § 4(B) is not a condition, but is a contractual covenant. *Cf. Sun Microsystems, Inc. v. Microsoft Corp*., 81 F. Supp. 2d 1026, 1031-32 (N.D. Cal. 2000) ("*Sun II*") (where Sun licensed Microsoft to create only derivative works compatible with other Sun software, Microsoft's "compatibility obligations" were covenants because the license was not specifically conditioned on their fulfillment).

To recover for copyright infringement based on breach of a license agreement, (1) the copying must exceed the scope of the defendant's license and (2) the copyright owner's complaint must be grounded in an exclusive right of copyright (e.g., unlawful reproduction or distribution). *See Storage Tech. Corp. v. Custom Hardware Eng'g & Consulting, Inc*., 421 F.3d 1307, 1315-16 (Fed. Cir. 2005). Contractual rights, however, can be much broader:

[C]onsider a license in which the copyright owner grants a person the right to make one and only one copy of a book with the caveat that the licensee may not read the last ten pages.

Obviously, a licensee who made a hundred copies of the book would be liable for copyright infringement because the copying would violate the Copyright Act's prohibition on reproduction and would exceed the scope of the license. Alternatively, if the licensee made a single copy of the book, but read the last ten pages, the only cause of action would be for breach of contract, because reading a book does not violate any right protected by copyright law.

Id. at 1316. Consistent with this approach, we have held that* *the potential for infringement exists only where the licensee's action (1) exceeds the license's scope (2) in a manner that implicates one of the licensor's exclusive statutory rights. *See, e.g., Sun I*, 118 F.3d at 1121-22 (remanding for infringement determination where defendant allegedly violated a license term regulating the creation of derivative works).

Here, ToU § 4 contains certain restrictions that are grounded in Blizzard's exclusive rights of copyright and other restrictions that are not. For instance, ToU § 4(D) forbids creation of derivative works based on WoW without Blizzard's consent. A player who violates this prohibition would exceed the scope of her license and violate one of Blizzard's exclusive rights under the Copyright Act. In contrast, ToU § 4©(ii) prohibits a player's disruption of another player's game experience. *Id*. A player might violate this prohibition while playing the game by harassing another player with unsolicited instant messages. Although this conduct may violate the contractual covenants with Blizzard, it would not violate any of Blizzard's exclusive rights of copyright. The anti-bot provisions at issue in this case, ToU § 4(B)(ii) and (iii), are similarly covenants rather than conditions. A Glider user violates the covenants with Blizzard, but does not thereby commit copyright infringement because Glider does not infringe any of Blizzard's exclusive rights. For instance, the use does not alter or copy WoW software.

Were we to hold otherwise, Blizzard — or any software copyright holder — could designate any disfavored conduct during software use as copyright infringement, by purporting to condition the license on the player's abstention from the disfavored conduct. The rationale would be that because the conduct occurs while the player's computer is copying the software code into RAM in order for it to run, the violation is copyright infringement. This would allow software copyright owners far greater rights than Congress has generally conferred on copyright owners.

We conclude that for a licensee's violation of a contract to constitute copyright infringement, there must be a nexus between the condition and the licensor's exclusive rights of copyright. Here, WoW players do not commit copyright infringement by using Glider in violation of the ToU. MDY is thus not liable for secondary copyright infringement, which requires the existence of direct copyright infringement. *Grokster*, 545 U.S. at 930.

It follows that because MDY does not infringe Blizzard's copyrights, we need not resolve MDY's contention that Blizzard commits copyright misuse. Copyright misuse is an equitable defense to copyright infringement, the contours of which are still being defined. *See Practice Mgmt. Info. Corp. v. Am. Med.*

Ass'n, 121 F.3d 516, 520 (9th Cir. 1997). The remedy for copyright misuse is to deny the copyright holder the right to enforce its copyright during the period of misuse. Since MDY does not infringe, we do not consider whether Blizzard committed copyright misuse.

We thus reverse the district court's grant of summary judgment to Blizzard on its secondary copyright infringement claims. Accordingly, we must also vacate the portion of the district court's permanent injunction that barred MDY and Donnelly from "infringing, or contributing to the infringement of, Blizzard's copyrights in WoW software."

[…]

Discussion

1. *MDY* provides the Ninth Circuit's test for determining whether a licensee's breach of a copyright license will give rise to claims for copyright infringement. See *MDY Indus., LLC v. Blizzard Entm't, Inc., 629 F.3d 928, 940-41 (9th Cir. 2010)* ("[T]he potential for infringement exists only where the licensee's action (1) exceeds the license's scope (2) in a manner that implicates one of the licensor's exclusive statutory rights. […] [F]or a licensee's violation of a contract to constitute copyright infringement, there must be a nexus between the condition and the licensor's exclusive rights of copyright.“).
2. Both the Ninth Circuit in *MDY* and the Federal Circuit in *Jacobsen* rely on the following holdings from *Sun I*: (1) a licensee must act outside the scope of a copyright license in

order to be liable for copyright infringement, and (2) a copyright license's terms must be distinguished as either covenants or limitations on the scope of the license. Are the Ninth Circuit and the Federal Circuit both reading *Sun I* the same way?

3. Are the *Jacobsen* "covenant vs. condition" and the *MDY* "nexus" tests in conflict? Is *MDY*'s expansion of the test to require a nexus between the breached condition and an exclusive right of copyright necessary in order to avoid absurd results? See *MDY Indus., LLC v. Blizzard Entm't, Inc., 629 F.3d 928, 941 (9th Cir. 2010)* ("Were we to hold otherwise, Blizzard — or any software copyright holder — could designate any disfavored conduct during software use as copyright infringement, by purporting to condition the license on the player's abstention from the disfavored conduct. The rationale would be that because the conduct occurs while the player's computer is copying the software code into RAM in order for it to run, the violation is copyright infringement. This would allow software copyright owners far greater rights than Congress has generally conferred on copyright owners.")

4. Does the Ninth Circuit's recognition of a "distinct nexus between payment and all commercial copyright licenses" break the *MDY* test? How does this reconcile with Jacobsen's acknowledgement of the economic benefit conferred by an open source licensing scheme? See *Jacobsen v. Katzer, 535 F.3d 1373, 1382 (Fed. Cir. 2008)* ("The clear language of the Artistic License creates conditions to protect the economic rights at issue in the granting of a public license. These conditions govern the rights to modify and distribute the computer

programs and files included in the downloadable software package. The attribution and modification transparency requirements directly serve to drive traffic to the open source incubation page and to inform downstream users of the project, which is a significant economic goal of the copyright holder that the law will enforce.“).

5. In Jacobsen, the Federal Circuit held that certain terms of the Artistic License should be regarded as conditions: the requirement that licensees “duplicate all of the original copyright notices and associated disclaimers,” and the requirement that modifiers of the software insert “prominent notices in each changed file stating how and when you changed that file.” Would a court applying *MDY’s* nexus test* *find that these conditions of the artistic license bear a nexus to the licensor’s exclusive rights of copyright? If so, which rights? See 17 U.S.C. § 106(2) (exclusive right to prepare derivative works); and see 17 U.S.C. §§ 1202 (b)(1) and -©(1) (prohibiting removal of copyright management information and defining copyright management information to comprise copyright notices).

6. What, aside from breach of contractual conditions, will constitute actions outside the scope of a copyright license? See *Storage Tech. Corp. v. Custom Hardware Eng’g & Consulting, Inc., 421 F.3d 1307, 1315-16 (Fed. Cir. 2005)* (”[C]onsider a license in which the copyright owner grants a person the right to make one and only one copy of a book with the caveat that the licensee may not read the last ten pages. Obviously, a licensee who made a hundred copies of the book would be liable for copyright infringement because the copying would violate the Copyright

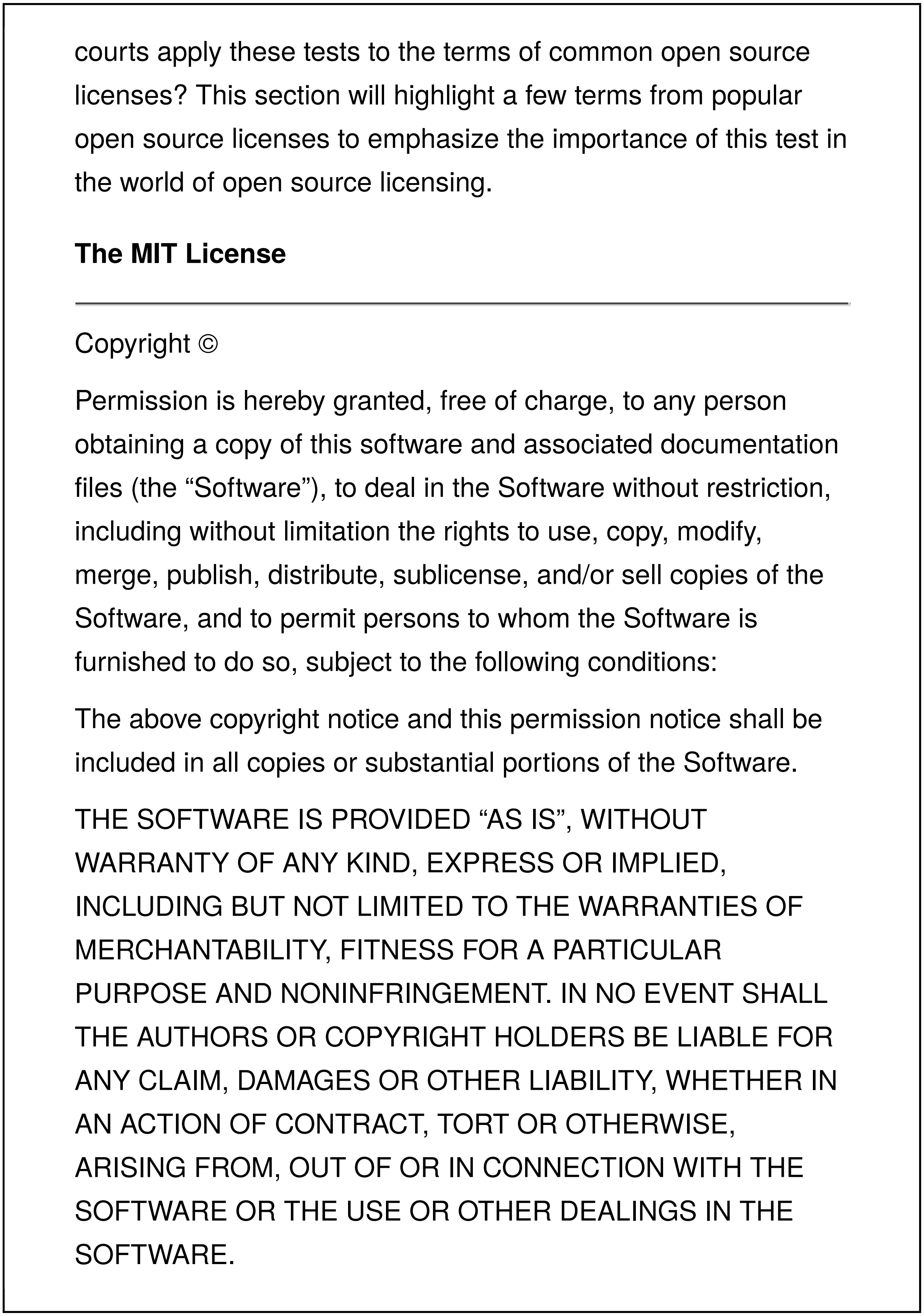
Act’s prohibition on reproduction and would exceed the scope of the license. Alternatively, if the licensee made a single copy of the book, but read the last ten pages, the only cause of action would be for breach of contract, because reading a book does not violate any right protected by copyright law.“).

What other kinds of complaints are grounded in an exclusive right of copyright? See* Adobe Sys. v. A & S Elecs., Inc., 153 F. Supp. 3d 1136, 1144 (N.D. Cal. 2015)* (finding that defendant exceeded scope of license and committed copyright infringement by selling CD keys - non-copyright protected 25-digit numbers - which facilitated the sale and use of software that defendant had no right to distribute).

7. Suppose a Licensor were to permit a Licensee to create derivative works of the Licensor’s program, provided that Licensee not sell any copies. What happens if Licensee owns a single, lawfully made copy of Licensor’s program and sells it on eBay? Does the first sale doctrine overcome the nexus between the breached condition and the Licensor’s exclusive rights of copyright if the exclusive right at issue is exhausted? See *Kirtsaeng v. John Wiley & Sons, Inc., 133 S. Ct. 1351, 1374 (2013)*; 17 U.S.C.S. § 109.

Common open source license terms: Covenants or Conditions?

Jacobsen and *MDY* provide possible frameworks for evaluating whether breach of a copyright license will be regarded as copyright infringement or breach of contract alone. How would



courts apply these tests to the terms of common open source licenses? This section will highlight a few terms from popular open source licenses to emphasize the importance of this test in the world of open source licensing.

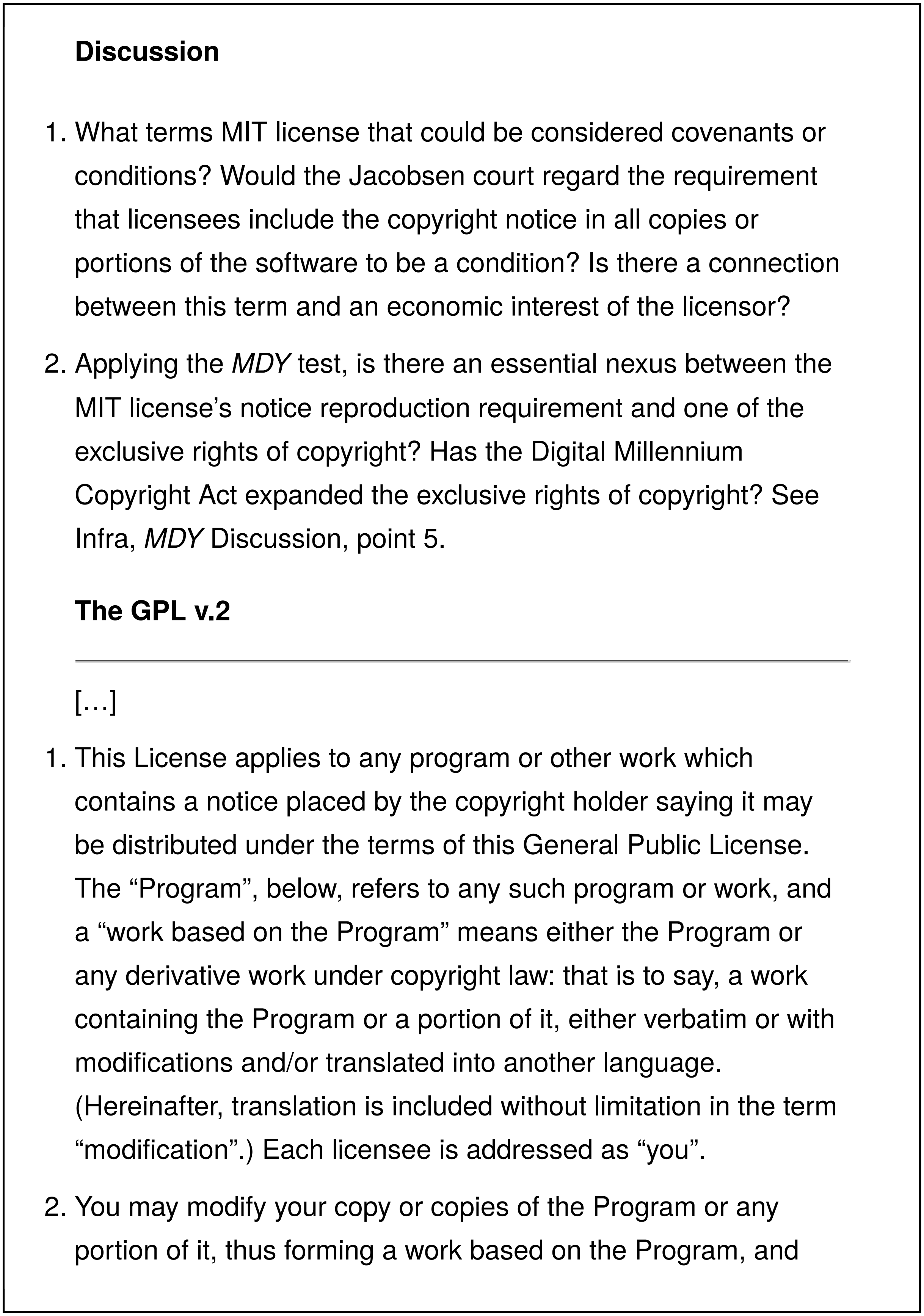
The MIT License

Copyright ©

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the “Software”), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED “AS IS”, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.



Discussion

1. What terms MIT license that could be considered covenants or conditions? Would the Jacobsen court regard the requirement that licensees include the copyright notice in all copies or portions of the software to be a condition? Is there a connection between this term and an economic interest of the licensor?
2. Applying the *MDY* test, is there an essential nexus between the MIT license's notice reproduction requirement and one of the exclusive rights of copyright? Has the Digital Millennium Copyright Act expanded the exclusive rights of copyright? See Infra, *MDY* Discussion, point 5.

The GPL v.2

[…]

1. This License applies to any program or other work which contains a notice placed by the copyright holder saying it may be distributed under the terms of this General Public License. The "Program", below, refers to any such program or work, and a "work based on the Program" means either the Program or any derivative work under copyright law: that is to say, a work containing the Program or a portion of it, either verbatim or with modifications and/or translated into another language. (Hereinafter, translation is included without limitation in the term "modification".) Each licensee is addressed as "you".
2. You may modify your copy or copies of the Program or any portion of it, thus forming a work based on the Program, and

> copy and distribute such modifications or work under the terms of Section 1 above, provided that you also meet all of these conditions:
>
> […]
>
> b) You must cause any work that you distribute or publish, that in whole or in part contains or is derived from the Program or any part thereof, to be licensed as a whole at no charge to all third parties under the terms of this License.
>
> These requirements apply to the modified work as a whole. If identifiable sections of that work are not derived from the Program, and can be reasonably considered independent and separate works in themselves, then this License, and its terms, do not apply to those sections when you distribute them as separate works. But when you distribute the same sections as part of a whole which is a work based on the Program, the distribution of the whole must be on the terms of this License, whose permissions for other licensees extend to the entire whole, and thus to each and every part regardless of who wrote it.
>
> 1. You may copy and distribute the Program (or a work based on it, under Section 2) in object code or executable form under the terms of Sections 1 and 2 above provided that you also do one of the following:
>
> a) Accompany it with the complete corresponding machine-readable source code, which must be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

b) Accompany it with a written offer, valid for at least three years, to give any third party, for a charge no more than your cost of physically performing source distribution, a complete machine-readable copy of the corresponding source code, to be distributed under the terms of Sections 1 and 2 above on a medium customarily used for software interchange; or,

c) Accompany it with the information you received as to the offer to distribute corresponding source code. (This alternative is allowed only for noncommercial distribution and only if you received the program in object code or executable form with such an offer, in accord with Subsection b above.)

The source code for a work means the preferred form of the work for making modifications to it. For an executable work, complete source code means all the source code for all modules it contains, plus any associated interface definition files, plus the scripts used to control compilation and installation of the executable.

[…]

Discussion

1. Is Section 2 requirement that independent and separate works, when combined with GPL v.2-licensed code to form a larger whole, must be licensed under the GPL v.2, a covenant or a condition? How about the Section 3 requirement that distribution of any GPL v.2-licensed code in binary form (as an executable program or as part of an executable program) be accompanied with distribution of the complete source code for the entire

binary? See* *Infra, *Jacobsen* Discussion, point 2.

2. Does the re-licensing requirement in Section 2 have a nexus to the licensor's exclusive rights of copyright? How about the source-mirroring requirement in Section 3? See *Versata Software, Inc. v. Ameriprise Fin., Inc., No. A-14-CA-12-SS, 2014 U.S. Dist. LEXIS 30934, at *14-15 (W.D. Tex. Mar. 10, 2014)* ("The 'viral' component of the GPL is separate and distinct from any copyright obligation. Copyright law imposes no open source obligations, and Ameriprise has not sued Versata for infringing XimpleWare's copyright by distributing VTD-XML without permission. Instead, Ameriprise has sued based on Versata's breach of an additional obligation: an affirmative promise to make its derivative work open source because it incorporated an open source program into its software. Ameriprise's claim therefore requires an 'extra element' in addition to reproduction or distribution: a failure to disclose the source code of the derivative software.“).

3. Do the GPLv2's relicensing terms govern derivative works alone, or is the requirement broader, extending to works that would not be considered derivative works under copyright law? See 7 Wash J.L. Tech. & Arts 265, 271 (2012) ("[T]he reference in Section 2(b) to a "work that in whole or in part contains … the Program," could be construed as including any work that incorporates code from the Program, no matter how insignificant and with no regard to whether the included code would be protectable under the Copyright Act.“). If the re-licensing requirement is intended to apply to works that are not derivative works, could the requirement still be a condition? Would it still

have a nexus to the licensor’s exclusive rights of copyright?

The GPL v.3

[…]

1. Protecting Users’ Legal Rights From Anti-Circumvention Law.

   No covered work shall be deemed part of an effective technological measure under any applicable law fulfilling obligations under article 11 of the WIPO copyright treaty adopted on 20 December 1996, or similar laws prohibiting or restricting circumvention of such measures.

   When you convey a covered work, you waive any legal power to forbid circumvention of technological measures to the extent such circumvention is effected by exercising rights under this License with respect to the covered work, and you disclaim any intention to limit operation or modification of the work as a means of enforcing, against the work’s users, your or third parties’ legal rights to forbid circumvention of technological measures. […]

1. Termination.

   You may not propagate or modify a covered work except as expressly provided under this License. Any attempt otherwise to propagate or modify it is void, and will automatically terminate your rights under this License (including any patent licenses granted under the third paragraph of section 11).

   However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, you do not qualify to receive new licenses for the same material under section 10.

Discussion

1. Read Section 3 of the GPLv3. How would a court apply the MDY and Jacobsen tests to a contractual condition bears a nexus to an exclusive right of copyright, but not to the licensor's exclusive rights of copyright in the copyrighted work being licensed? Could a licensor's expectation interest in the scheme of Section 3 be fulfilled without specific performance?
2. Section 8 of the GPLv3 includes a remedy scheme that permanently reinstates a license if a licensee meets the criteria for curing the violation. Does the mere existence of this provision require treating all of the GPL's terms as covenants

rather than conditions? *Sun Microsystems, Inc. v. Microsoft Corp., 81 F. Supp. 2d 1026, 1033 (N.D. Cal. 2000)* ("If Sun could sue for copyright infringement immediately upon Microsoft's failure to fully meet the compatibility requirements, the remedies scheme would be frustrated and Microsoft would not get the full benefit of its bargained for cure periods.“).

Notes

Part II

Peace and Its Dividends

Chapter 3

Join the GPL Cooperation Commitment

gplcc.github.io

Join the GPL Cooperation Commitment

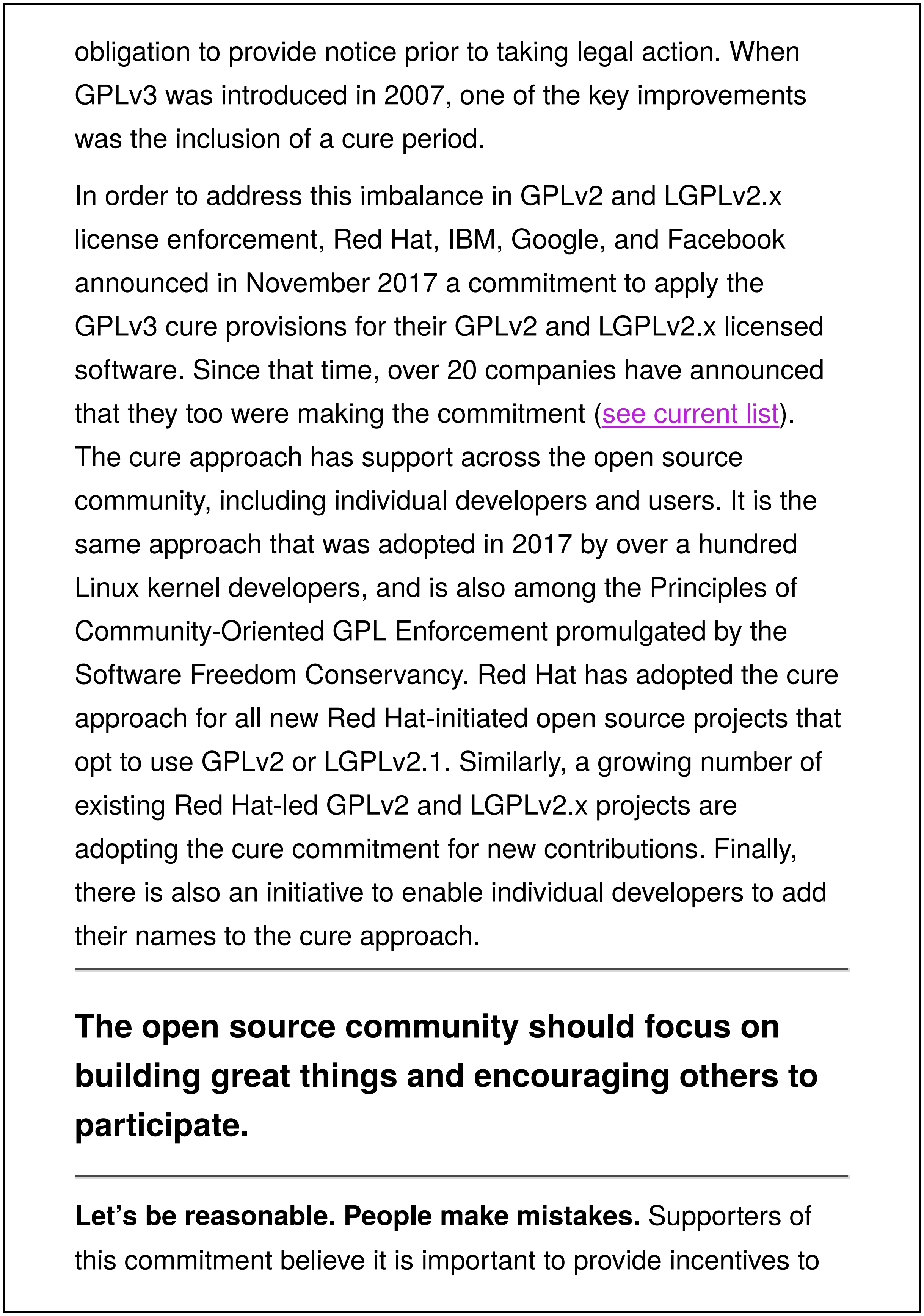
12-15 minutes

Join with leading companies, developers, and other leaders in the open source community who have all committed to provide GPLv2 and LGPLv2.x licensees a fair chance to correct violations before their licenses are terminated.

Our goal is to reduce opportunities for abusive enforcement tactics and, more broadly, to promote greater predictability in the enforcement of GPLv2 and LGPLv2.x licenses. Through this initiative, we hope ultimately to increase participation in the use and development of open source software by helping to ensure that enforcement, when it takes place, is fair and predictable.

The GPL Cooperation Commitment is a statement by GPLv2 and LGPLv2.x copyright holders and other supporters that gives licensees a fair chance to correct violations before their licenses are terminated.

The “automatic termination” feature of GPLv2 and LGPLv2.x does not provide an express “cure” period in the event of a violation. This means that a single act of inadvertent noncompliance could give rise to an infringement claim, with no

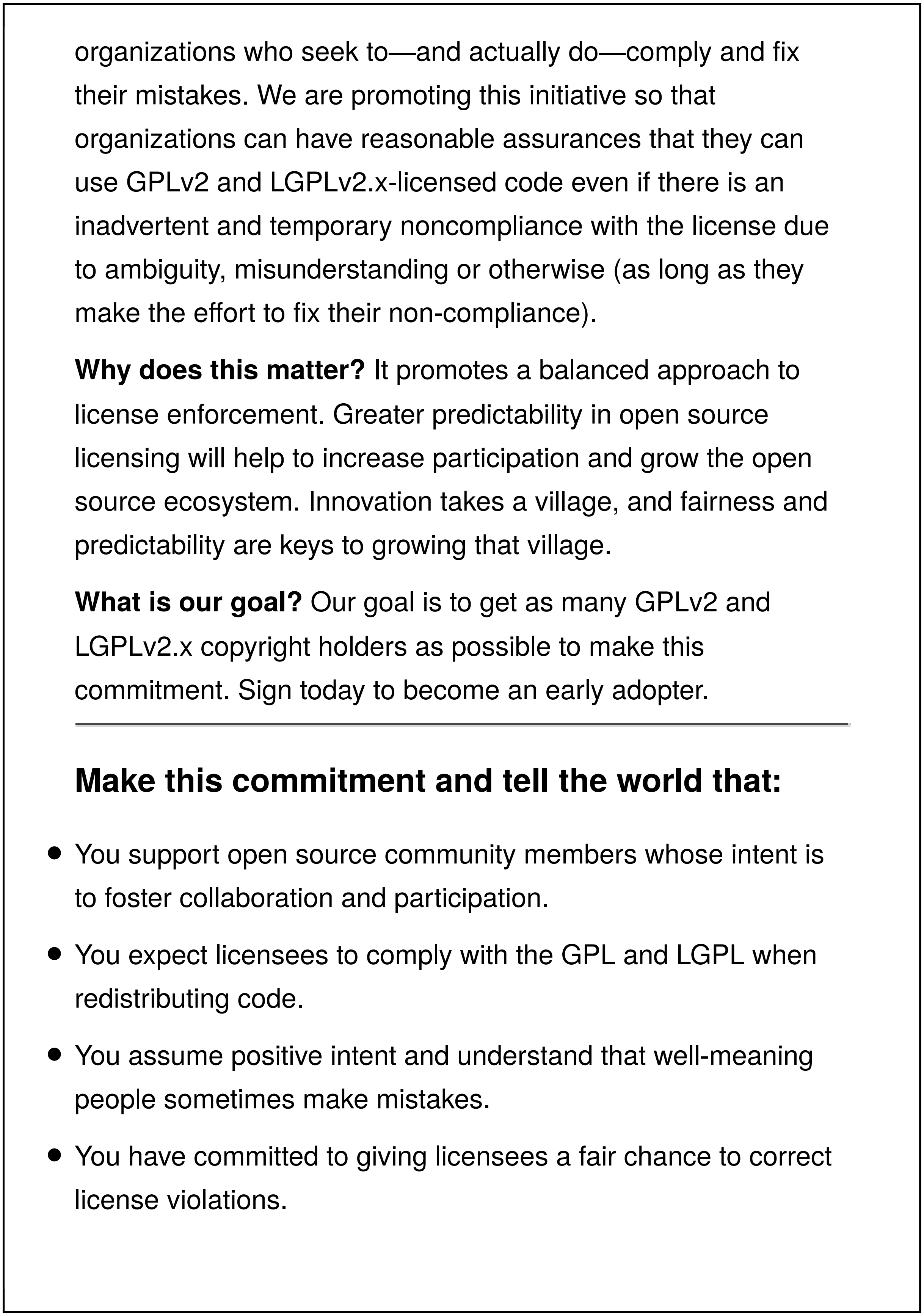


obligation to provide notice prior to taking legal action. When GPLv3 was introduced in 2007, one of the key improvements was the inclusion of a cure period.

In order to address this imbalance in GPLv2 and LGPLv2.x license enforcement, Red Hat, IBM, Google, and Facebook announced in November 2017 a commitment to apply the GPLv3 cure provisions for their GPLv2 and LGPLv2.x licensed software. Since that time, over 20 companies have announced that they too were making the commitment (see current list). The cure approach has support across the open source community, including individual developers and users. It is the same approach that was adopted in 2017 by over a hundred Linux kernel developers, and is also among the Principles of Community-Oriented GPL Enforcement promulgated by the Software Freedom Conservancy. Red Hat has adopted the cure approach for all new Red Hat-initiated open source projects that opt to use GPLv2 or LGPLv2.1. Similarly, a growing number of existing Red Hat-led GPLv2 and LGPLv2.x projects are adopting the cure commitment for new contributions. Finally, there is also an initiative to enable individual developers to add their names to the cure approach.

The open source community should focus on building great things and encouraging others to participate.

Let’s be reasonable. People make mistakes. Supporters of this commitment believe it is important to provide incentives to



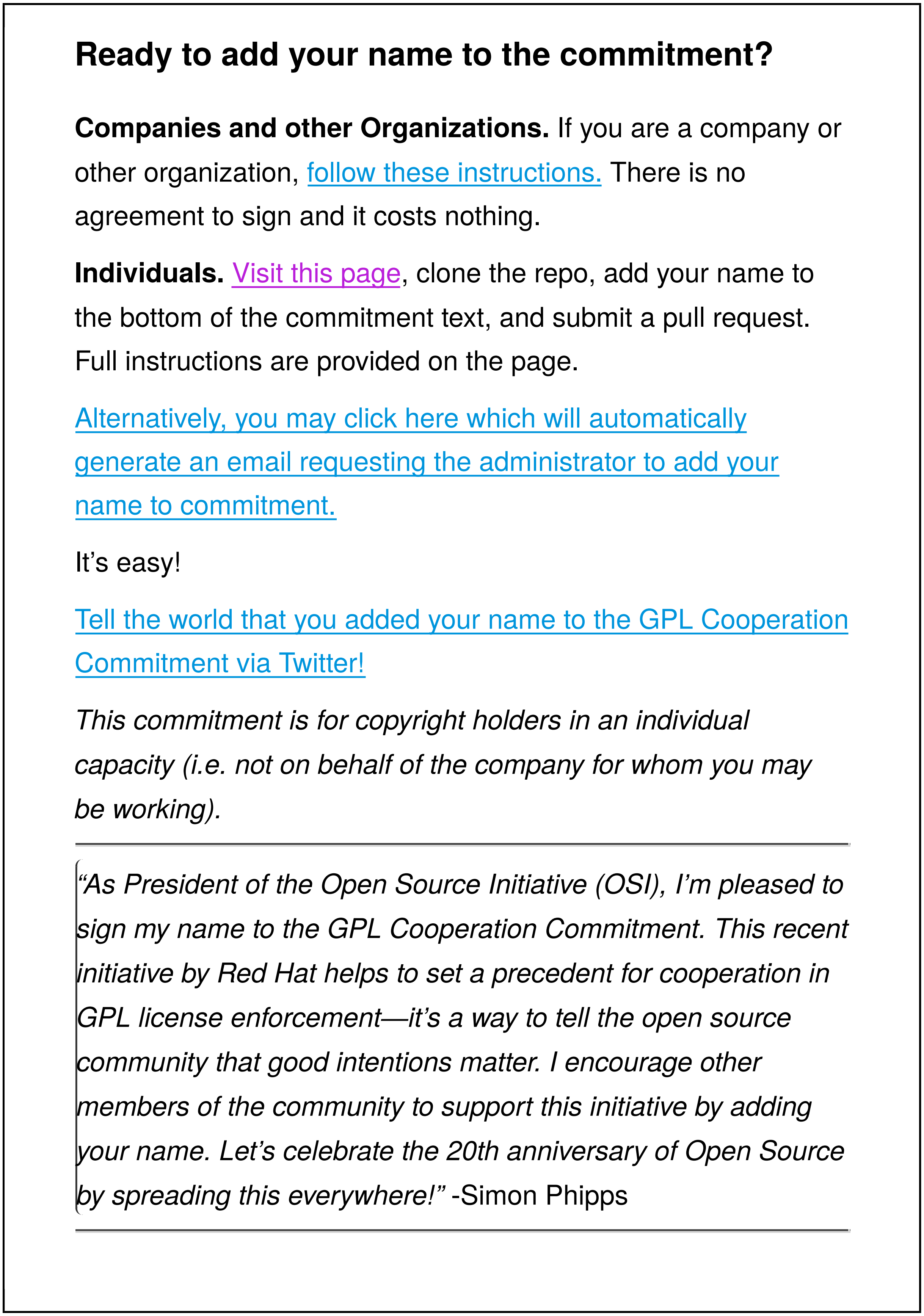
organizations who seek to—and actually do—comply and fix their mistakes. We are promoting this initiative so that organizations can have reasonable assurances that they can use GPLv2 and LGPLv2.x-licensed code even if there is an inadvertent and temporary noncompliance with the license due to ambiguity, misunderstanding or otherwise (as long as they make the effort to fix their non-compliance).

Why does this matter? It promotes a balanced approach to license enforcement. Greater predictability in open source licensing will help to increase participation and grow the open source ecosystem. Innovation takes a village, and fairness and predictability are keys to growing that village.

What is our goal? Our goal is to get as many GPLv2 and LGPLv2.x copyright holders as possible to make this commitment. Sign today to become an early adopter.

Make this commitment and tell the world that:

- You support open source community members whose intent is to foster collaboration and participation.
- You expect licensees to comply with the GPL and LGPL when redistributing code.
- You assume positive intent and understand that well-meaning people sometimes make mistakes.
- You have committed to giving licensees a fair chance to correct license violations.



Ready to add your name to the commitment?

Companies and other Organizations. If you are a company or other organization, follow these instructions. There is no agreement to sign and it costs nothing.

Individuals. Visit this page, clone the repo, add your name to the bottom of the commitment text, and submit a pull request. Full instructions are provided on the page.

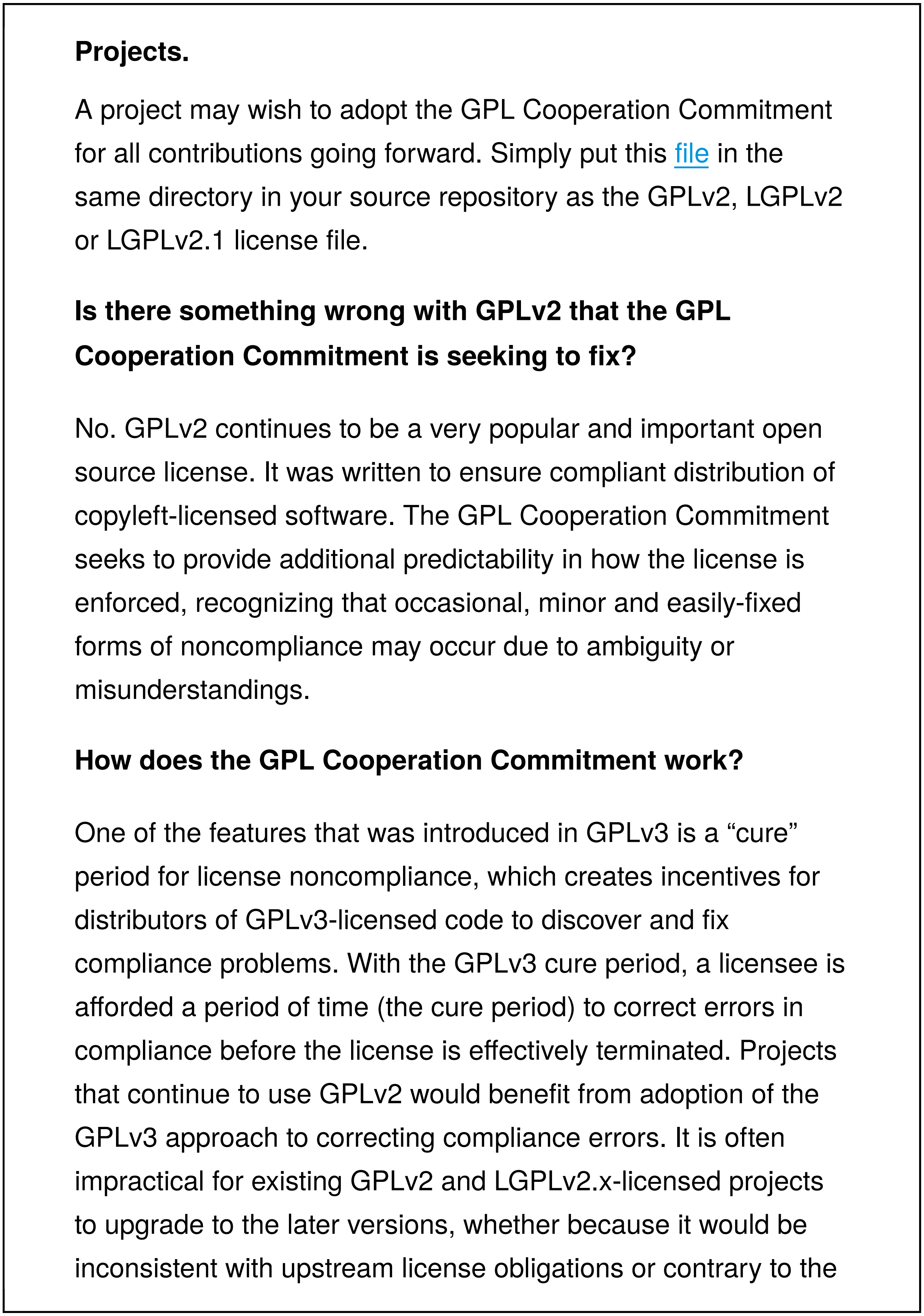
Alternatively, you may click here which will automatically generate an email requesting the administrator to add your name to commitment.

It's easy!

Tell the world that you added your name to the GPL Cooperation Commitment via Twitter!

This commitment is for copyright holders in an individual capacity (i.e. not on behalf of the company for whom you may be working).

> *"As President of the Open Source Initiative (OSI), I'm pleased to sign my name to the GPL Cooperation Commitment. This recent initiative by Red Hat helps to set a precedent for cooperation in GPL license enforcement—it's a way to tell the open source community that good intentions matter. I encourage other members of the community to support this initiative by adding your name. Let's celebrate the 20th anniversary of Open Source by spreading this everywhere!"* -Simon Phipps



Projects.

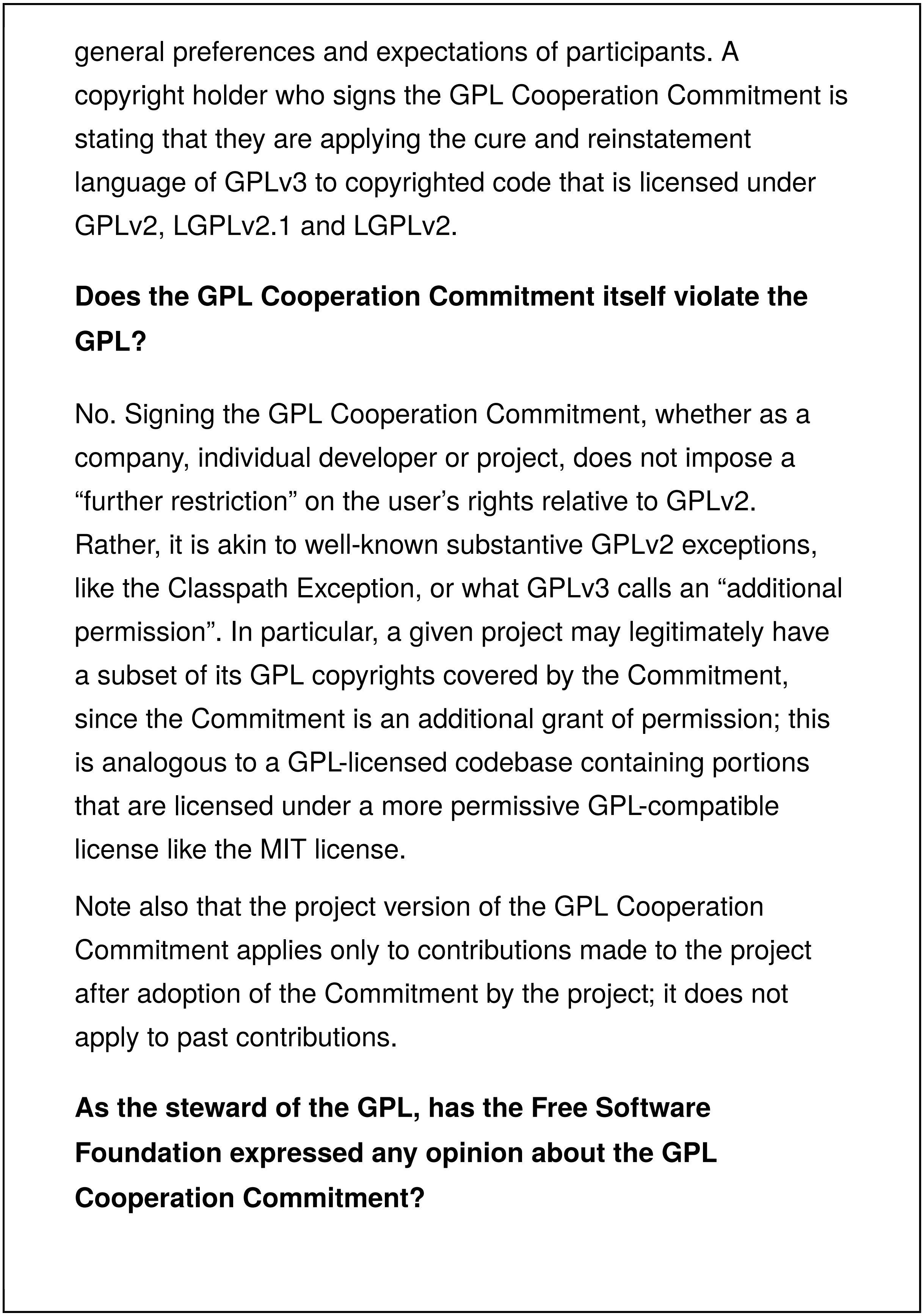
A project may wish to adopt the GPL Cooperation Commitment for all contributions going forward. Simply put this file in the same directory in your source repository as the GPLv2, LGPLv2 or LGPLv2.1 license file.

Is there something wrong with GPLv2 that the GPL Cooperation Commitment is seeking to fix?

No. GPLv2 continues to be a very popular and important open source license. It was written to ensure compliant distribution of copyleft-licensed software. The GPL Cooperation Commitment seeks to provide additional predictability in how the license is enforced, recognizing that occasional, minor and easily-fixed forms of noncompliance may occur due to ambiguity or misunderstandings.

How does the GPL Cooperation Commitment work?

One of the features that was introduced in GPLv3 is a “cure” period for license noncompliance, which creates incentives for distributors of GPLv3-licensed code to discover and fix compliance problems. With the GPLv3 cure period, a licensee is afforded a period of time (the cure period) to correct errors in compliance before the license is effectively terminated. Projects that continue to use GPLv2 would benefit from adoption of the GPLv3 approach to correcting compliance errors. It is often impractical for existing GPLv2 and LGPLv2.x-licensed projects to upgrade to the later versions, whether because it would be inconsistent with upstream license obligations or contrary to the



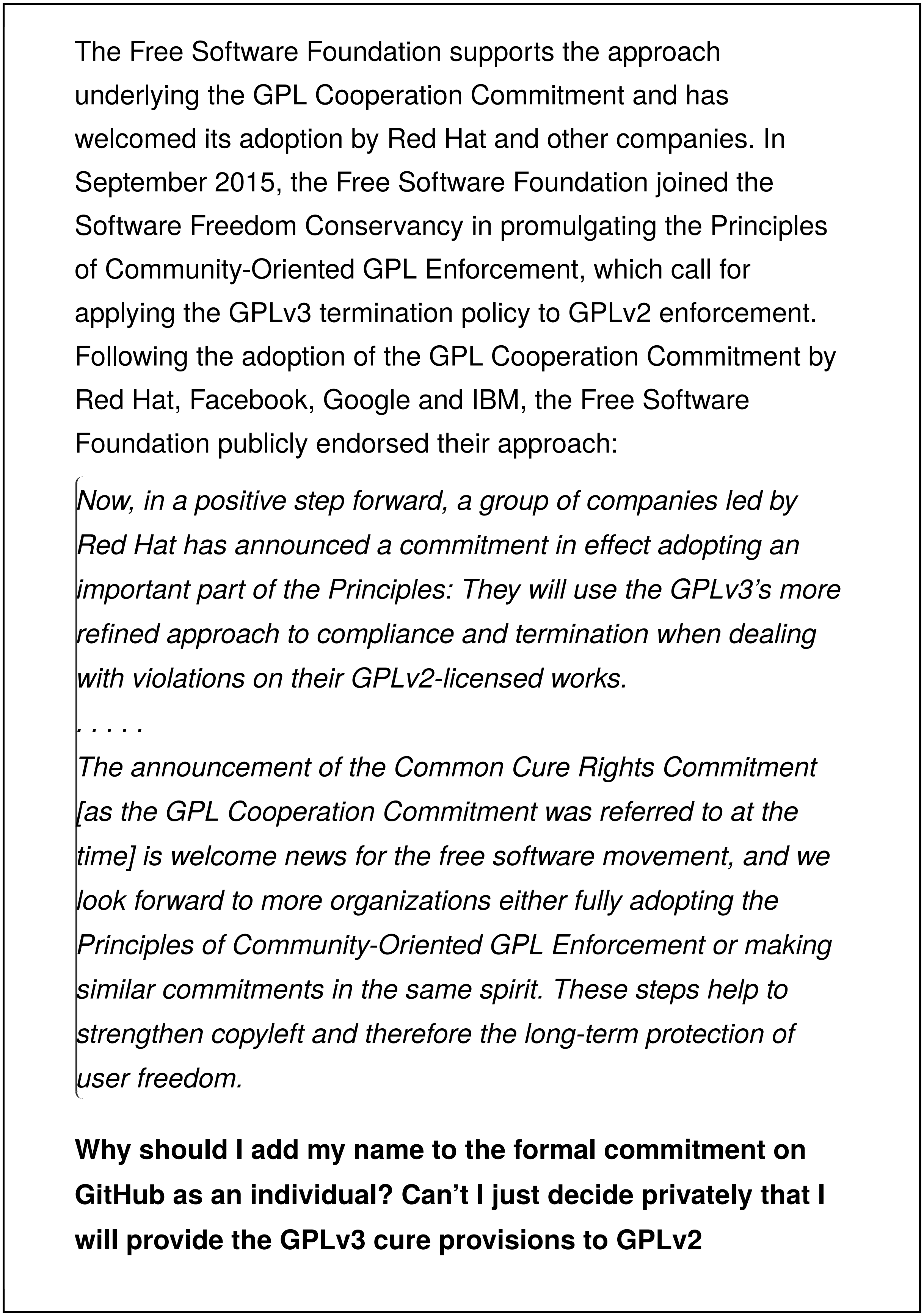
general preferences and expectations of participants. A copyright holder who signs the GPL Cooperation Commitment is stating that they are applying the cure and reinstatement language of GPLv3 to copyrighted code that is licensed under GPLv2, LGPLv2.1 and LGPLv2.

Does the GPL Cooperation Commitment itself violate the GPL?

No. Signing the GPL Cooperation Commitment, whether as a company, individual developer or project, does not impose a “further restriction” on the user’s rights relative to GPLv2. Rather, it is akin to well-known substantive GPLv2 exceptions, like the Classpath Exception, or what GPLv3 calls an “additional permission”. In particular, a given project may legitimately have a subset of its GPL copyrights covered by the Commitment, since the Commitment is an additional grant of permission; this is analogous to a GPL-licensed codebase containing portions that are licensed under a more permissive GPL-compatible license like the MIT license.

Note also that the project version of the GPL Cooperation Commitment applies only to contributions made to the project after adoption of the Commitment by the project; it does not apply to past contributions.

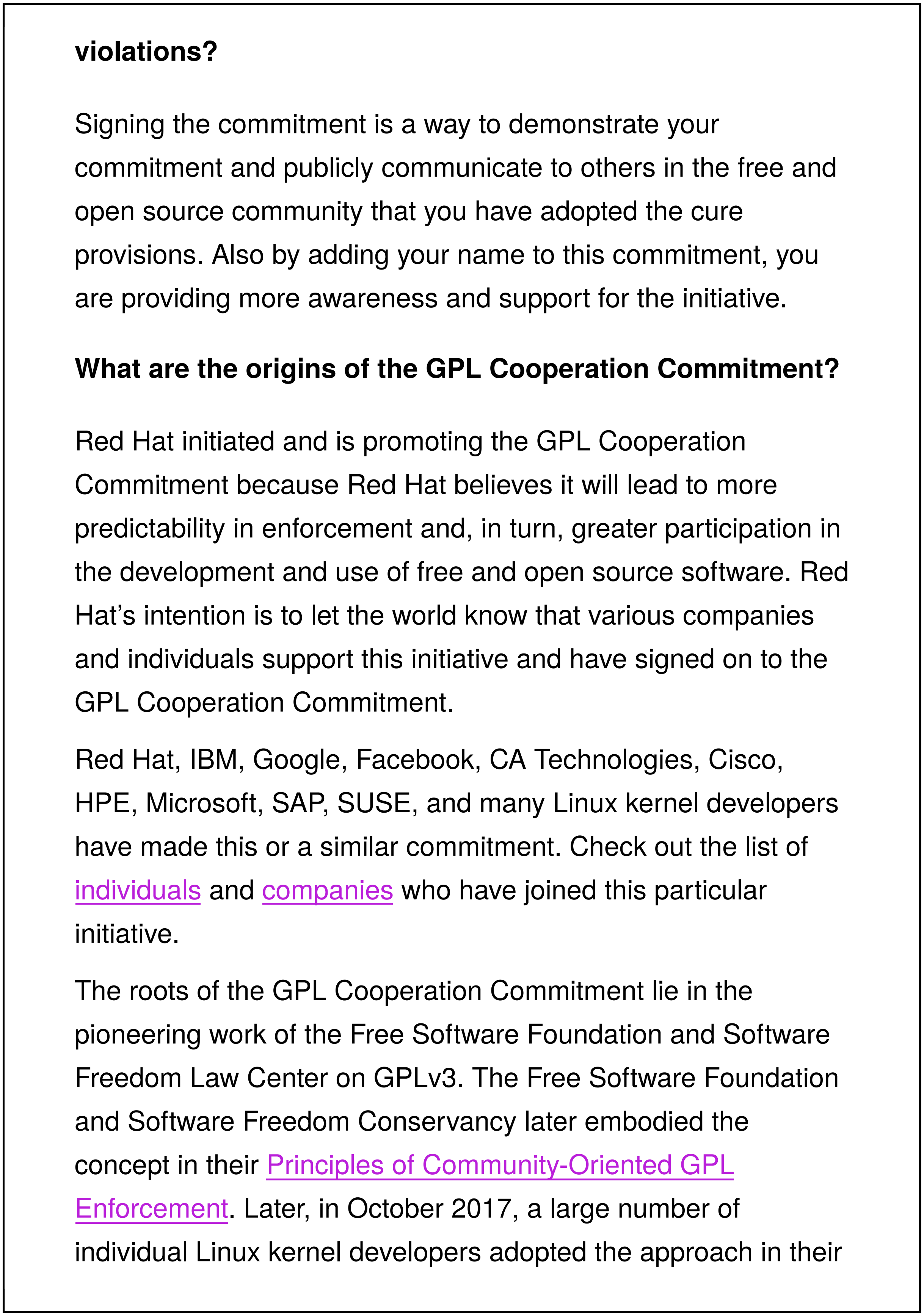
As the steward of the GPL, has the Free Software Foundation expressed any opinion about the GPL Cooperation Commitment?



The Free Software Foundation supports the approach underlying the GPL Cooperation Commitment and has welcomed its adoption by Red Hat and other companies. In September 2015, the Free Software Foundation joined the Software Freedom Conservancy in promulgating the Principles of Community-Oriented GPL Enforcement, which call for applying the GPLv3 termination policy to GPLv2 enforcement. Following the adoption of the GPL Cooperation Commitment by Red Hat, Facebook, Google and IBM, the Free Software Foundation publicly endorsed their approach:

> *Now, in a positive step forward, a group of companies led by Red Hat has announced a commitment in effect adopting an important part of the Principles: They will use the GPLv3's more refined approach to compliance and termination when dealing with violations on their GPLv2-licensed works.*
>
> *. . . . .*
>
> *The announcement of the Common Cure Rights Commitment [as the GPL Cooperation Commitment was referred to at the time] is welcome news for the free software movement, and we look forward to more organizations either fully adopting the Principles of Community-Oriented GPL Enforcement or making similar commitments in the same spirit. These steps help to strengthen copyleft and therefore the long-term protection of user freedom.*

Why should I add my name to the formal commitment on GitHub as an individual? Can't I just decide privately that I will provide the GPLv3 cure provisions to GPLv2



violations?

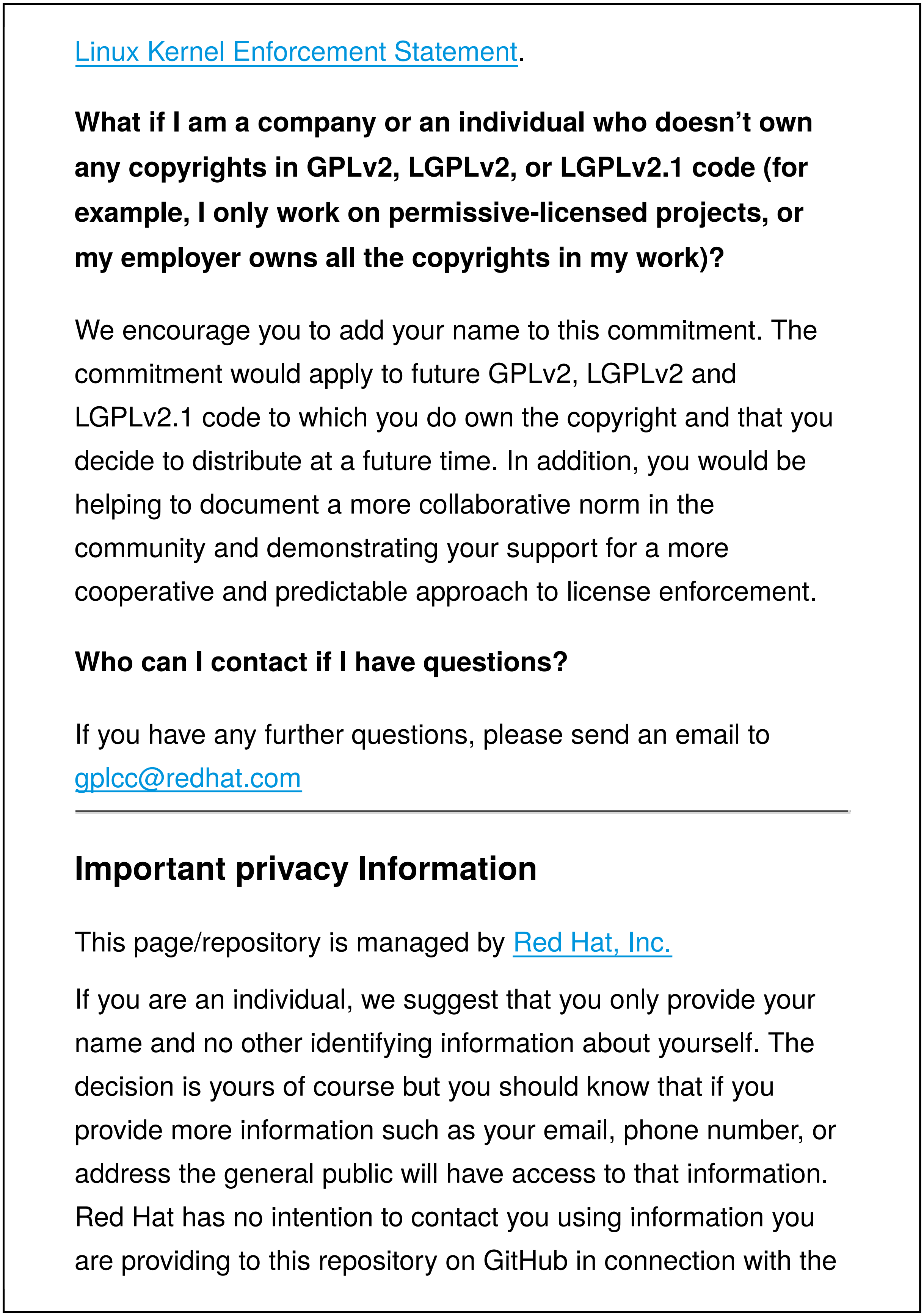
Signing the commitment is a way to demonstrate your commitment and publicly communicate to others in the free and open source community that you have adopted the cure provisions. Also by adding your name to this commitment, you are providing more awareness and support for the initiative.

What are the origins of the GPL Cooperation Commitment?

Red Hat initiated and is promoting the GPL Cooperation Commitment because Red Hat believes it will lead to more predictability in enforcement and, in turn, greater participation in the development and use of free and open source software. Red Hat's intention is to let the world know that various companies and individuals support this initiative and have signed on to the GPL Cooperation Commitment.

Red Hat, IBM, Google, Facebook, CA Technologies, Cisco, HPE, Microsoft, SAP, SUSE, and many Linux kernel developers have made this or a similar commitment. Check out the list of individuals and companies who have joined this particular initiative.

The roots of the GPL Cooperation Commitment lie in the pioneering work of the Free Software Foundation and Software Freedom Law Center on GPLv3. The Free Software Foundation and Software Freedom Conservancy later embodied the concept in their Principles of Community-Oriented GPL Enforcement. Later, in October 2017, a large number of individual Linux kernel developers adopted the approach in their



Linux Kernel Enforcement Statement.

What if I am a company or an individual who doesn't own any copyrights in GPLv2, LGPLv2, or LGPLv2.1 code (for example, I only work on permissive-licensed projects, or my employer owns all the copyrights in my work)?

We encourage you to add your name to this commitment. The commitment would apply to future GPLv2, LGPLv2 and LGPLv2.1 code to which you do own the copyright and that you decide to distribute at a future time. In addition, you would be helping to document a more collaborative norm in the community and demonstrating your support for a more cooperative and predictable approach to license enforcement.

Who can I contact if I have questions?

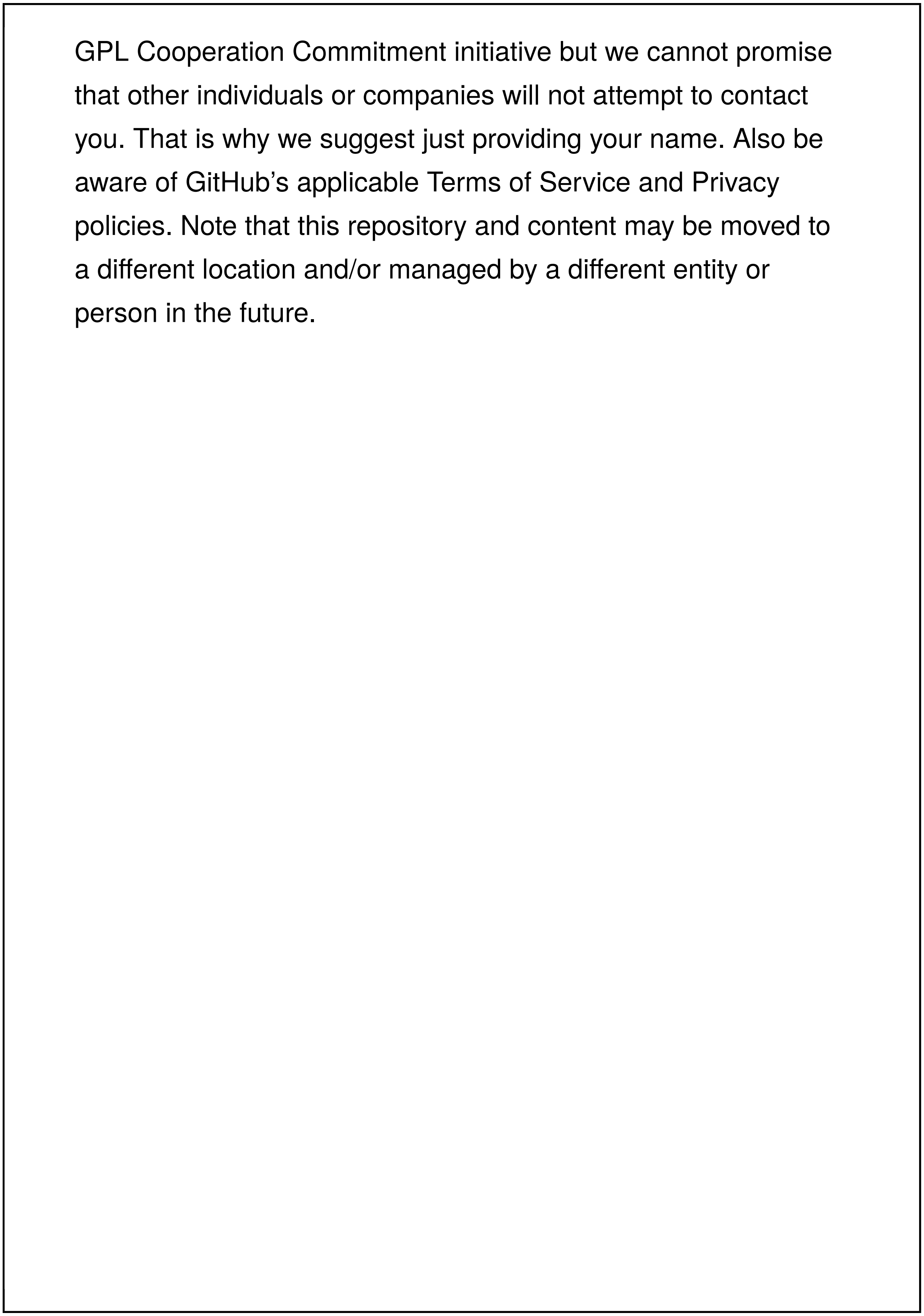
If you have any further questions, please send an email to gplcc@redhat.com

---

Important privacy Information

This page/repository is managed by Red Hat, Inc.

If you are an individual, we suggest that you only provide your name and no other identifying information about yourself. The decision is yours of course but you should know that if you provide more information such as your email, phone number, or address the general public will have access to that information. Red Hat has no intention to contact you using information you are providing to this repository on GitHub in connection with the



GPL Cooperation Commitment initiative but we cannot promise that other individuals or companies will not attempt to contact you. That is why we suggest just providing your name. Also be aware of GitHub’s applicable Terms of Service and Privacy policies. Note that this repository and content may be moved to a different location and/or managed by a different entity or person in the future.

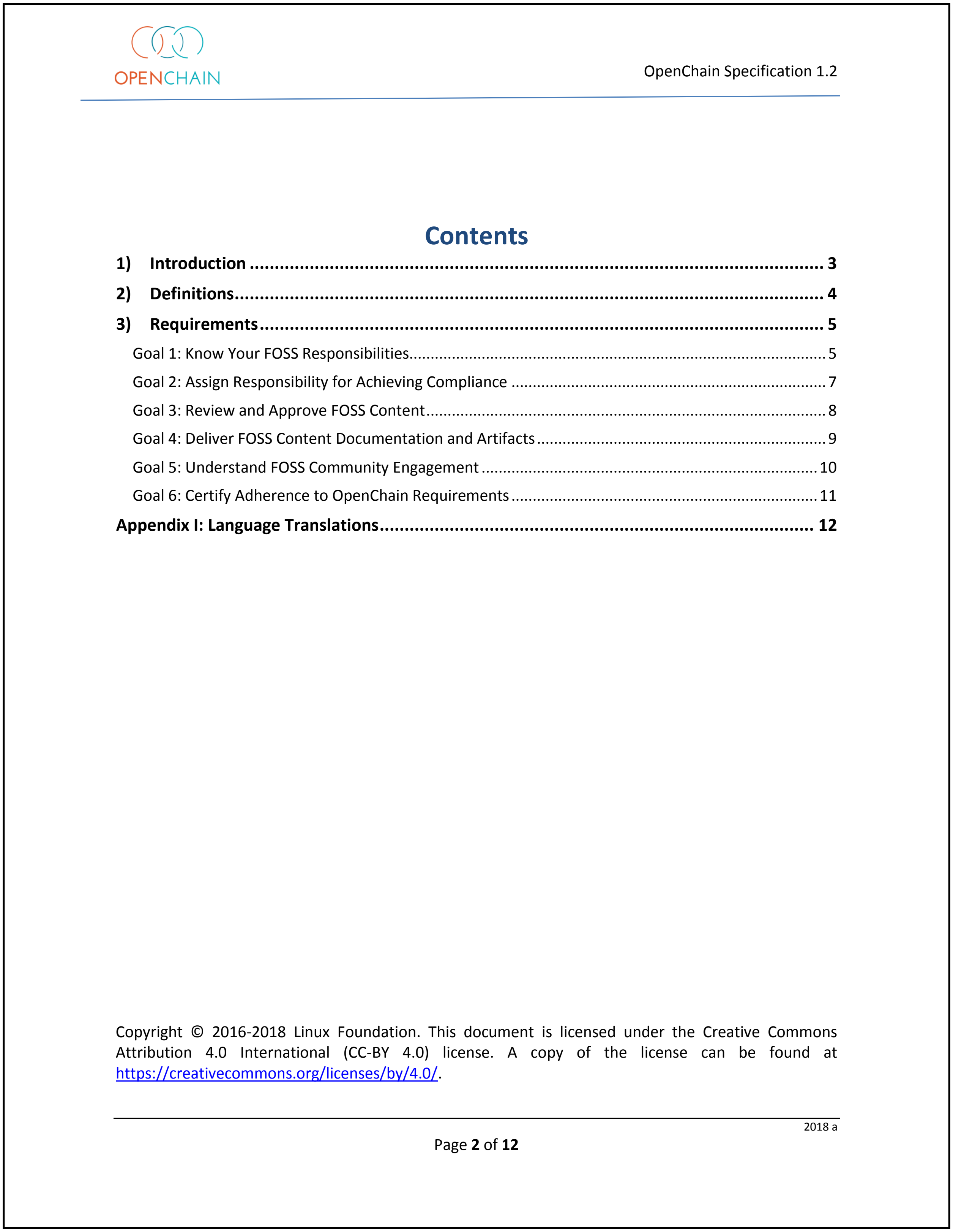
Chapter 4

OpenChain Specification, Version 1.2

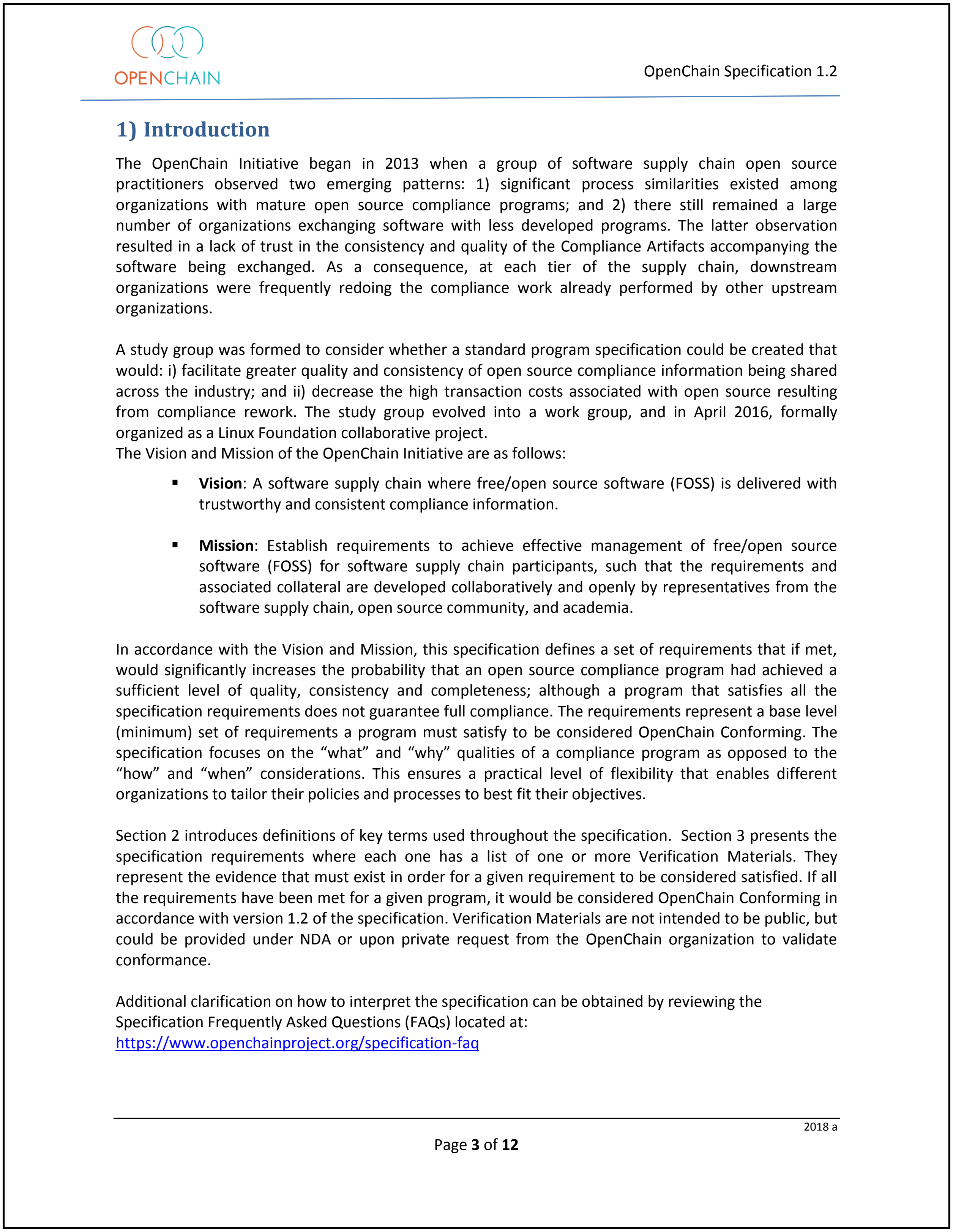
OPENCHAIN

OpenChain Specification

Version 1.2



Contents



1) Introduction

The OpenChain Initiative began in 2013 when a group of software supply chain open source practitioners observed two emerging patterns: 1) significant process similarities existed among organizations with mature open source compliance programs; and 2) there still remained a large number of organizations exchanging software with less developed programs. The latter observation resulted in a lack of trust in the consistency and quality of the Compliance Artifacts accompanying the software being exchanged. As a consequence, at each tier of the supply chain, downstream organizations were frequently redoing the compliance work already performed by other upstream organizations.

A study group was formed to consider whether a standard program specification could be created that would: i) facilitate greater quality and consistency of open source compliance information being shared across the industry; and ii) decrease the high transaction costs associated with open source resulting from compliance rework. The study group evolved into a work group, and in April 2016, formally organized as a Linux Foundation collaborative project.
The Vision and Mission of the OpenChain Initiative are as follows:

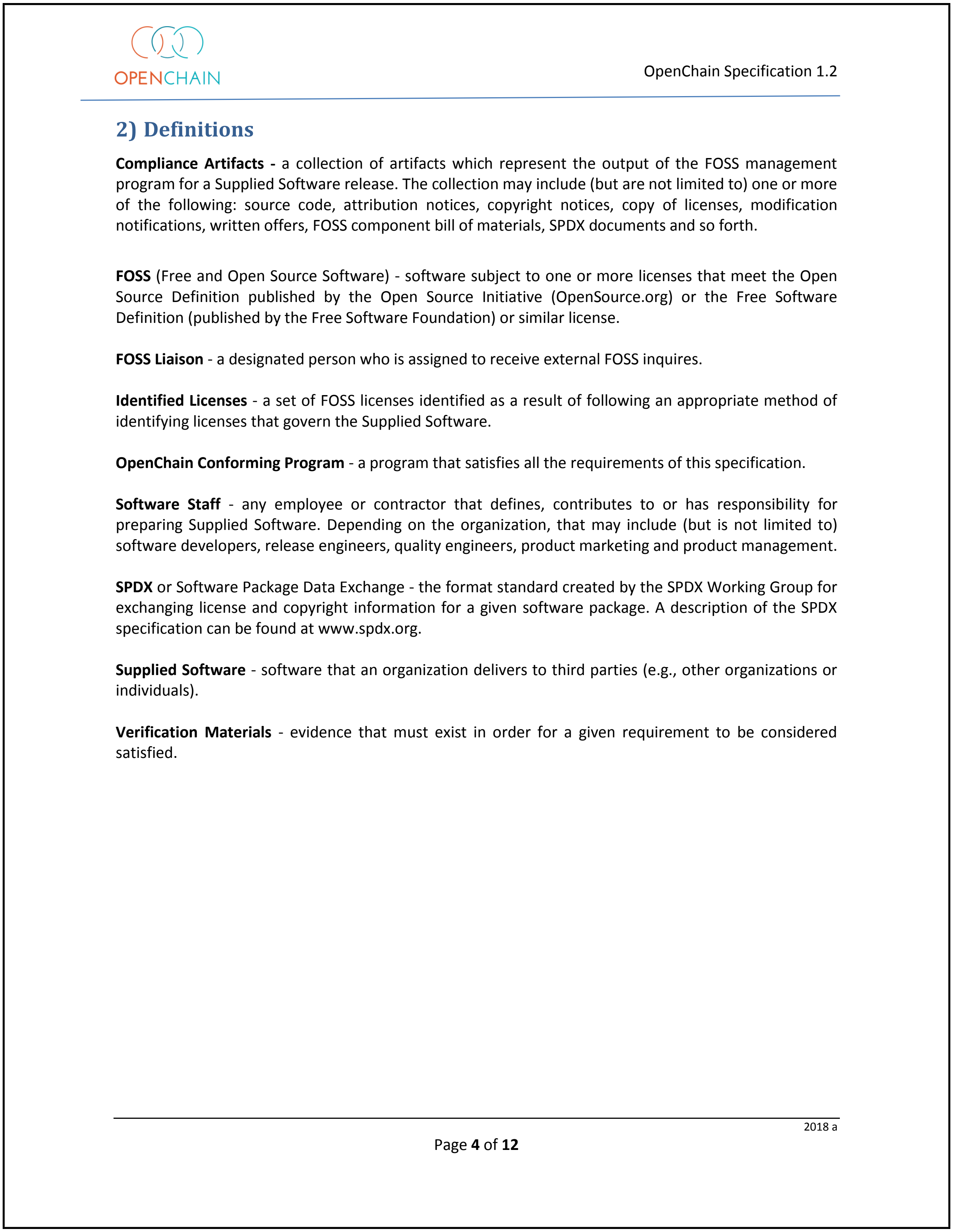
- **Vision**: A software supply chain where free/open source software (FOSS) is delivered with trustworthy and consistent compliance information.

- **Mission**: Establish requirements to achieve effective management of free/open source software (FOSS) for software supply chain participants, such that the requirements and associated collateral are developed collaboratively and openly by representatives from the software supply chain, open source community, and academia.

In accordance with the Vision and Mission, this specification defines a set of requirements that if met, would significantly increases the probability that an open source compliance program had achieved a sufficient level of quality, consistency and completeness; although a program that satisfies all the specification requirements does not guarantee full compliance. The requirements represent a base level (minimum) set of requirements a program must satisfy to be considered OpenChain Conforming. The specification focuses on the “what” and “why” qualities of a compliance program as opposed to the “how” and “when” considerations. This ensures a practical level of flexibility that enables different organizations to tailor their policies and processes to best fit their objectives.

Section 2 introduces definitions of key terms used throughout the specification. Section 3 presents the specification requirements where each one has a list of one or more Verification Materials. They represent the evidence that must exist in order for a given requirement to be considered satisfied. If all the requirements have been met for a given program, it would be considered OpenChain Conforming in accordance with version 1.2 of the specification. Verification Materials are not intended to be public, but could be provided under NDA or upon private request from the OpenChain organization to validate conformance.

Additional clarification on how to interpret the specification can be obtained by reviewing the Specification Frequently Asked Questions (FAQs) located at:
https://www.openchainproject.org/specification-faq



2) Definitions

Compliance Artifacts - a collection of artifacts which represent the output of the FOSS management program for a Supplied Software release. The collection may include (but are not limited to) one or more of the following: source code, attribution notices, copyright notices, copy of licenses, modification notifications, written offers, FOSS component bill of materials, SPDX documents and so forth.

FOSS (Free and Open Source Software) - software subject to one or more licenses that meet the Open Source Definition published by the Open Source Initiative (OpenSource.org) or the Free Software Definition (published by the Free Software Foundation) or similar license.

FOSS Liaison - a designated person who is assigned to receive external FOSS inquires.

Identified Licenses - a set of FOSS licenses identified as a result of following an appropriate method of identifying licenses that govern the Supplied Software.

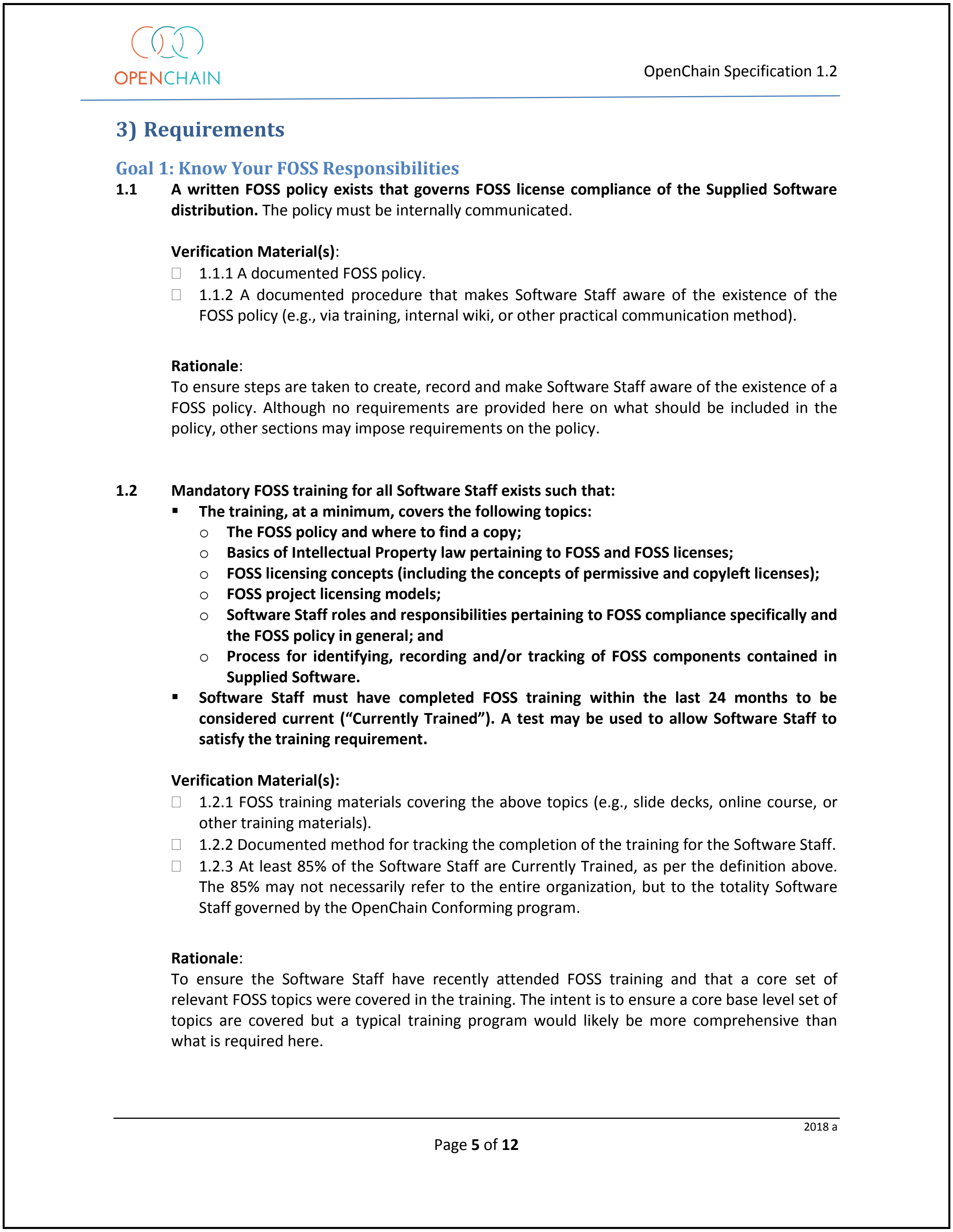
OpenChain Conforming Program - a program that satisfies all the requirements of this specification.

Software Staff - any employee or contractor that defines, contributes to or has responsibility for preparing Supplied Software. Depending on the organization, that may include (but is not limited to) software developers, release engineers, quality engineers, product marketing and product management.

SPDX or Software Package Data Exchange - the format standard created by the SPDX Working Group for exchanging license and copyright information for a given software package. A description of the SPDX specification can be found at www.spdx.org.

Supplied Software - software that an organization delivers to third parties (e.g., other organizations or individuals).

Verification Materials - evidence that must exist in order for a given requirement to be considered satisfied.



3) Requirements

Goal 1: Know Your FOSS Responsibilities

1.1 **A written FOSS policy exists that governs FOSS license compliance of the Supplied Software distribution.** The policy must be internally communicated.

Verification Material(s):

□ 1.1.1 A documented FOSS policy.

□ 1.1.2 A documented procedure that makes Software Staff aware of the existence of the FOSS policy (e.g., via training, internal wiki, or other practical communication method).

Rationale:
To ensure steps are taken to create, record and make Software Staff aware of the existence of a FOSS policy. Although no requirements are provided here on what should be included in the policy, other sections may impose requirements on the policy.

1.2 **Mandatory FOSS training for all Software Staff exists such that:**

- **The training, at a minimum, covers the following topics:**
  - **The FOSS policy and where to find a copy;**
  - **Basics of Intellectual Property law pertaining to FOSS and FOSS licenses;**
  - **FOSS licensing concepts (including the concepts of permissive and copyleft licenses);**
  - **FOSS project licensing models;**
  - **Software Staff roles and responsibilities pertaining to FOSS compliance specifically and the FOSS policy in general; and**
  - **Process for identifying, recording and/or tracking of FOSS components contained in Supplied Software.**
- **Software Staff must have completed FOSS training within the last 24 months to be considered current ("Currently Trained"). A test may be used to allow Software Staff to satisfy the training requirement.**

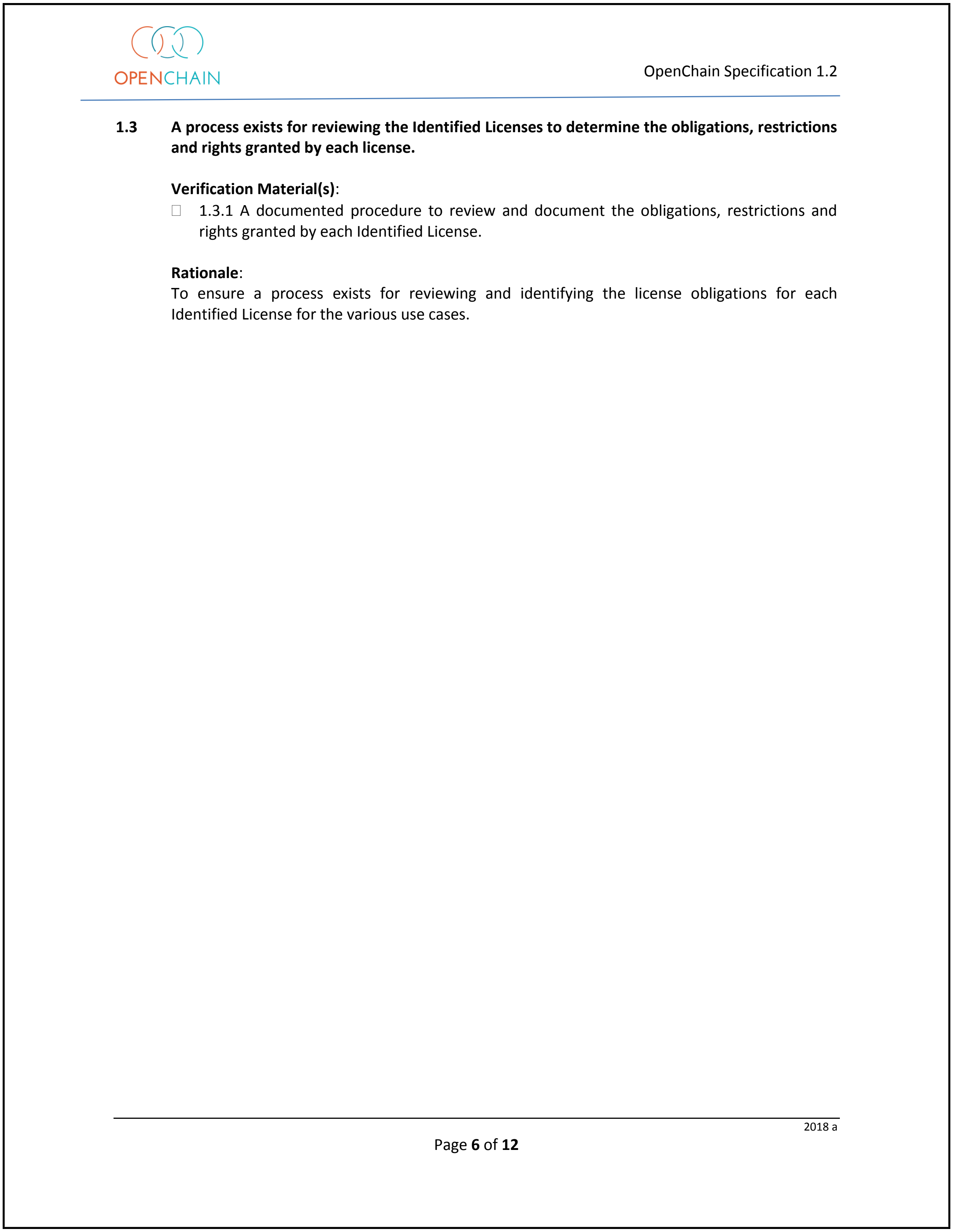
Verification Material(s):

□ 1.2.1 FOSS training materials covering the above topics (e.g., slide decks, online course, or other training materials).

□ 1.2.2 Documented method for tracking the completion of the training for the Software Staff.

□ 1.2.3 At least 85% of the Software Staff are Currently Trained, as per the definition above. The 85% may not necessarily refer to the entire organization, but to the totality Software Staff governed by the OpenChain Conforming program.

Rationale:
To ensure the Software Staff have recently attended FOSS training and that a core set of relevant FOSS topics were covered in the training. The intent is to ensure a core base level set of topics are covered but a typical training program would likely be more comprehensive than what is required here.



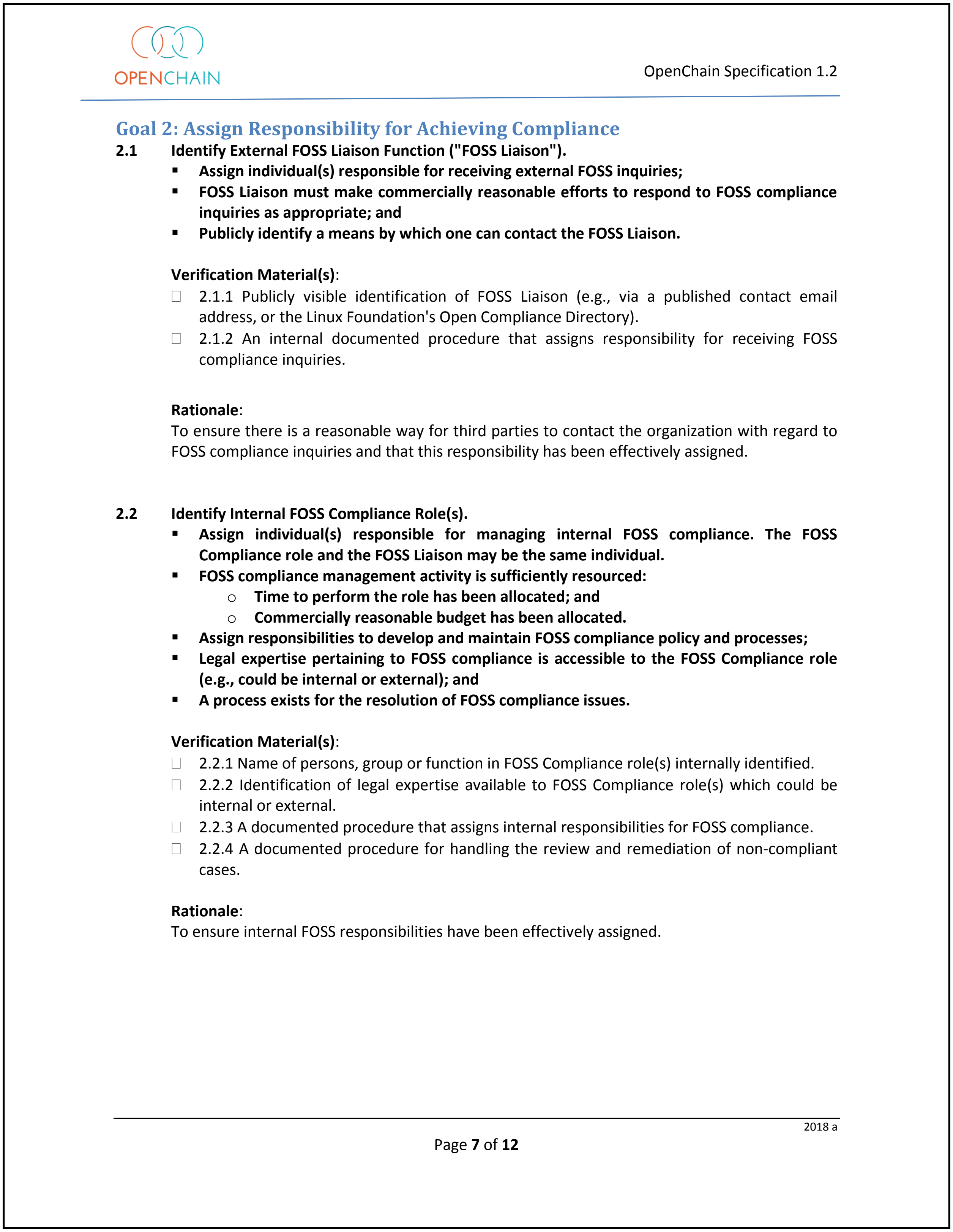
1.3 A process exists for reviewing the Identified Licenses to determine the obligations, restrictions and rights granted by each license.

Verification Material(s):

□ 1.3.1 A documented procedure to review and document the obligations, restrictions and rights granted by each Identified License.

Rationale:

To ensure a process exists for reviewing and identifying the license obligations for each Identified License for the various use cases.



Goal 2: Assign Responsibility for Achieving Compliance

2.1 Identify External FOSS Liaison Function ("FOSS Liaison").

- **Assign individual(s) responsible for receiving external FOSS inquiries;**
- **FOSS Liaison must make commercially reasonable efforts to respond to FOSS compliance inquiries as appropriate; and**
- **Publicly identify a means by which one can contact the FOSS Liaison.**

Verification Material(s):

□ 2.1.1 Publicly visible identification of FOSS Liaison (e.g., via a published contact email address, or the Linux Foundation's Open Compliance Directory).

□ 2.1.2 An internal documented procedure that assigns responsibility for receiving FOSS compliance inquiries.

Rationale:

To ensure there is a reasonable way for third parties to contact the organization with regard to FOSS compliance inquiries and that this responsibility has been effectively assigned.

2.2 Identify Internal FOSS Compliance Role(s).

- **Assign individual(s) responsible for managing internal FOSS compliance. The FOSS Compliance role and the FOSS Liaison may be the same individual.**
- **FOSS compliance management activity is sufficiently resourced:**
  - **Time to perform the role has been allocated; and**
  - **Commercially reasonable budget has been allocated.**
- **Assign responsibilities to develop and maintain FOSS compliance policy and processes;**
- **Legal expertise pertaining to FOSS compliance is accessible to the FOSS Compliance role (e.g., could be internal or external); and**
- **A process exists for the resolution of FOSS compliance issues.**

Verification Material(s):

□ 2.2.1 Name of persons, group or function in FOSS Compliance role(s) internally identified.

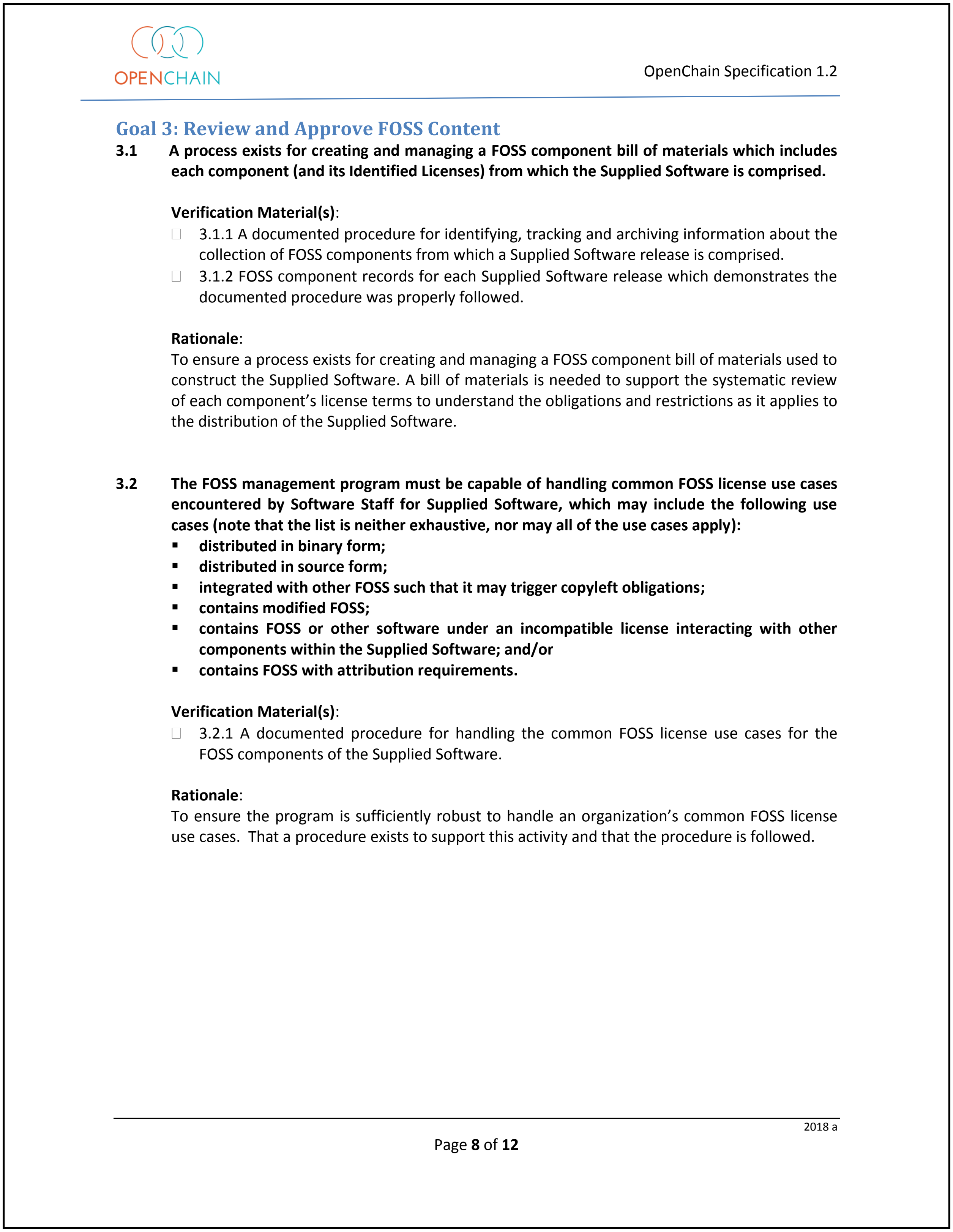
□ 2.2.2 Identification of legal expertise available to FOSS Compliance role(s) which could be internal or external.

□ 2.2.3 A documented procedure that assigns internal responsibilities for FOSS compliance.

□ 2.2.4 A documented procedure for handling the review and remediation of non-compliant cases.

Rationale:

To ensure internal FOSS responsibilities have been effectively assigned.



Goal 3: Review and Approve FOSS Content

3.1 A process exists for creating and managing a FOSS component bill of materials which includes each component (and its Identified Licenses) from which the Supplied Software is comprised.

Verification Material(s):

□ 3.1.1 A documented procedure for identifying, tracking and archiving information about the collection of FOSS components from which a Supplied Software release is comprised.

□ 3.1.2 FOSS component records for each Supplied Software release which demonstrates the documented procedure was properly followed.

Rationale:

To ensure a process exists for creating and managing a FOSS component bill of materials used to construct the Supplied Software. A bill of materials is needed to support the systematic review of each component's license terms to understand the obligations and restrictions as it applies to the distribution of the Supplied Software.

3.2 The FOSS management program must be capable of handling common FOSS license use cases encountered by Software Staff for Supplied Software, which may include the following use cases (note that the list is neither exhaustive, nor may all of the use cases apply):

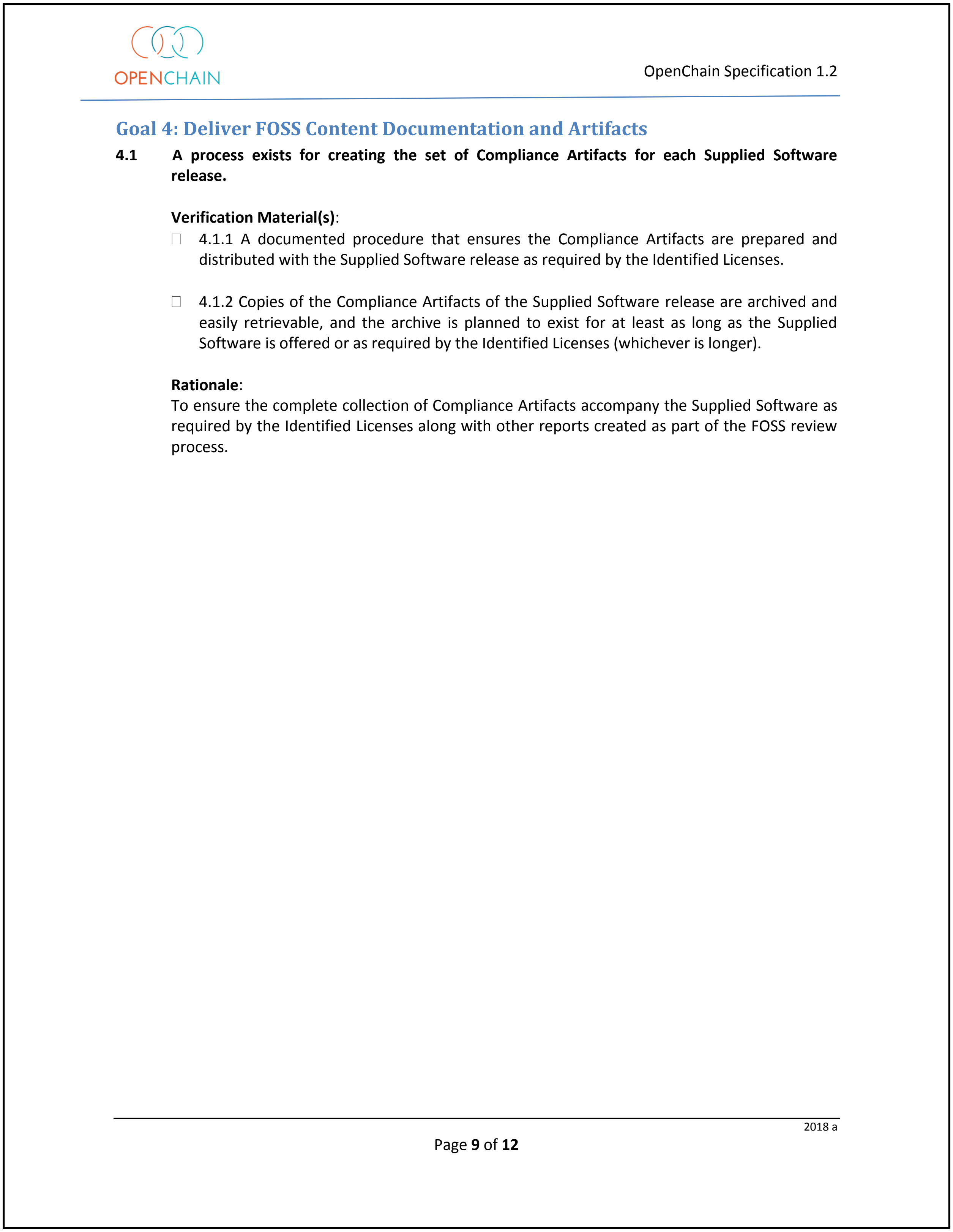
- **distributed in binary form;**
- **distributed in source form;**
- **integrated with other FOSS such that it may trigger copyleft obligations;**
- **contains modified FOSS;**
- **contains FOSS or other software under an incompatible license interacting with other components within the Supplied Software; and/or**
- **contains FOSS with attribution requirements.**

Verification Material(s):

□ 3.2.1 A documented procedure for handling the common FOSS license use cases for the FOSS components of the Supplied Software.

Rationale:

To ensure the program is sufficiently robust to handle an organization's common FOSS license use cases. That a procedure exists to support this activity and that the procedure is followed.



OpenChain Specification 1.2

Goal 4: Deliver FOSS Content Documentation and Artifacts

4.1 A process exists for creating the set of Compliance Artifacts for each Supplied Software release.

Verification Material(s):

□ 4.1.1 A documented procedure that ensures the Compliance Artifacts are prepared and distributed with the Supplied Software release as required by the Identified Licenses.

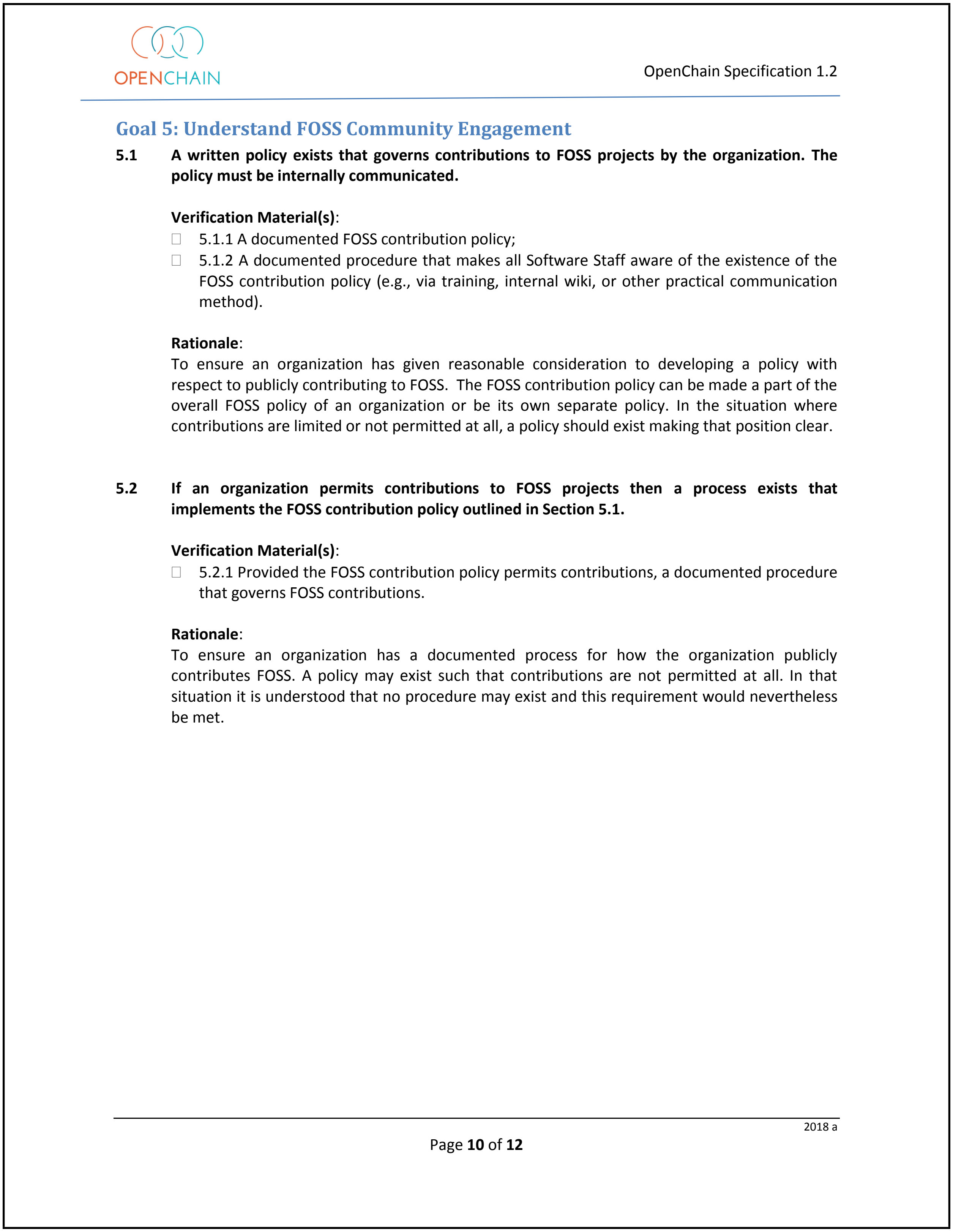
□ 4.1.2 Copies of the Compliance Artifacts of the Supplied Software release are archived and easily retrievable, and the archive is planned to exist for at least as long as the Supplied Software is offered or as required by the Identified Licenses (whichever is longer).

Rationale:

To ensure the complete collection of Compliance Artifacts accompany the Supplied Software as required by the Identified Licenses along with other reports created as part of the FOSS review process.

2018 a

Page **9** of **12**



Goal 5: Understand FOSS Community Engagement

5.1 A written policy exists that governs contributions to FOSS projects by the organization. The policy must be internally communicated.

Verification Material(s):

□ 5.1.1 A documented FOSS contribution policy;

□ 5.1.2 A documented procedure that makes all Software Staff aware of the existence of the FOSS contribution policy (e.g., via training, internal wiki, or other practical communication method).

Rationale:

To ensure an organization has given reasonable consideration to developing a policy with respect to publicly contributing to FOSS. The FOSS contribution policy can be made a part of the overall FOSS policy of an organization or be its own separate policy. In the situation where contributions are limited or not permitted at all, a policy should exist making that position clear.

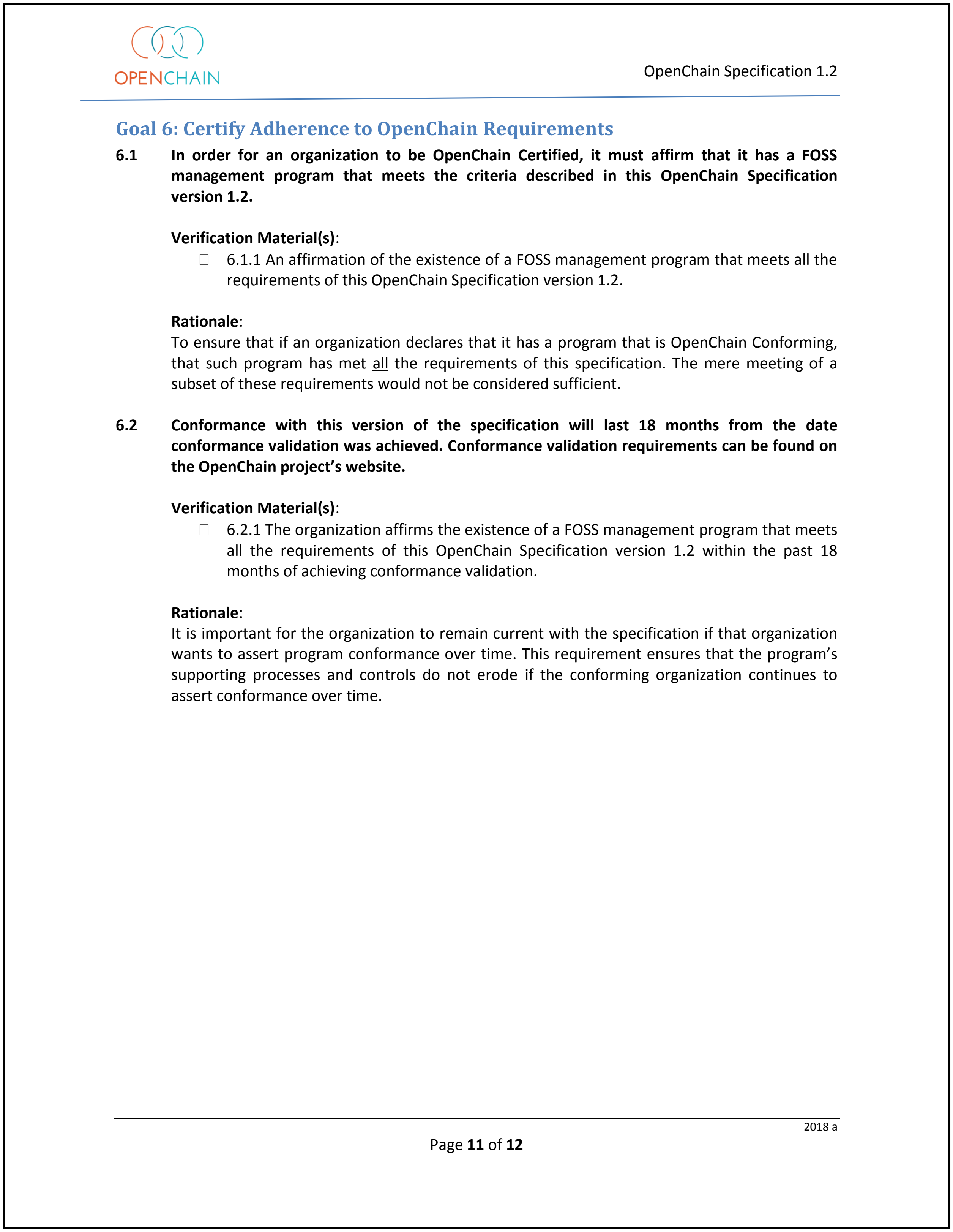
5.2 If an organization permits contributions to FOSS projects then a process exists that implements the FOSS contribution policy outlined in Section 5.1.

Verification Material(s):

□ 5.2.1 Provided the FOSS contribution policy permits contributions, a documented procedure that governs FOSS contributions.

Rationale:

To ensure an organization has a documented process for how the organization publicly contributes FOSS. A policy may exist such that contributions are not permitted at all. In that situation it is understood that no procedure may exist and this requirement would nevertheless be met.



Goal 6: Certify Adherence to OpenChain Requirements

6.1 In order for an organization to be OpenChain Certified, it must affirm that it has a FOSS management program that meets the criteria described in this OpenChain Specification version 1.2.

Verification Material(s):

□ 6.1.1 An affirmation of the existence of a FOSS management program that meets all the requirements of this OpenChain Specification version 1.2.

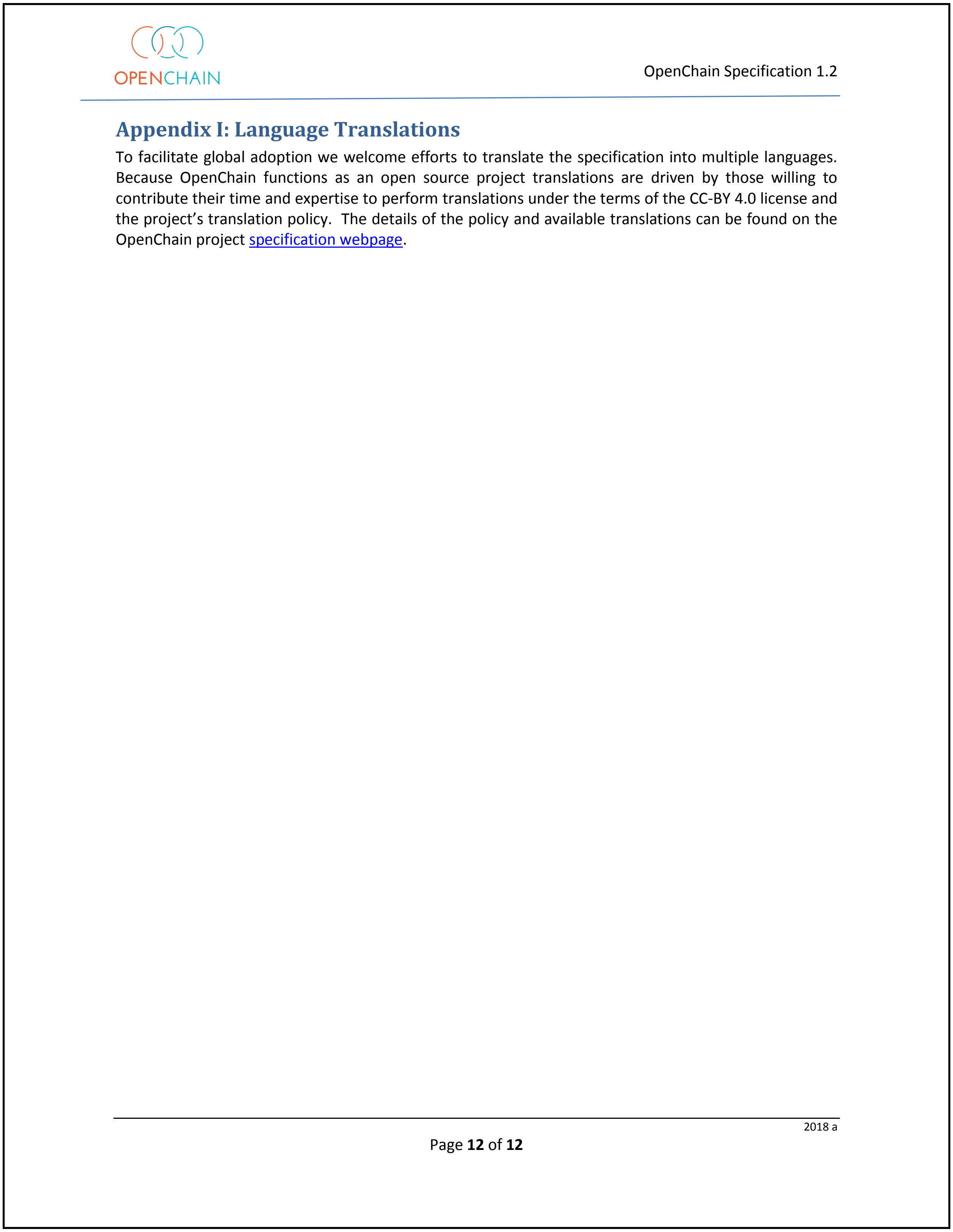
Rationale:
To ensure that if an organization declares that it has a program that is OpenChain Conforming, that such program has met all the requirements of this specification. The mere meeting of a subset of these requirements would not be considered sufficient.

6.2 Conformance with this version of the specification will last 18 months from the date conformance validation was achieved. Conformance validation requirements can be found on the OpenChain project's website.

Verification Material(s):

□ 6.2.1 The organization affirms the existence of a FOSS management program that meets all the requirements of this OpenChain Specification version 1.2 within the past 18 months of achieving conformance validation.

Rationale:
It is important for the organization to remain current with the specification if that organization wants to assert program conformance over time. This requirement ensures that the program's supporting processes and controls do not erode if the conforming organization continues to assert conformance over time.

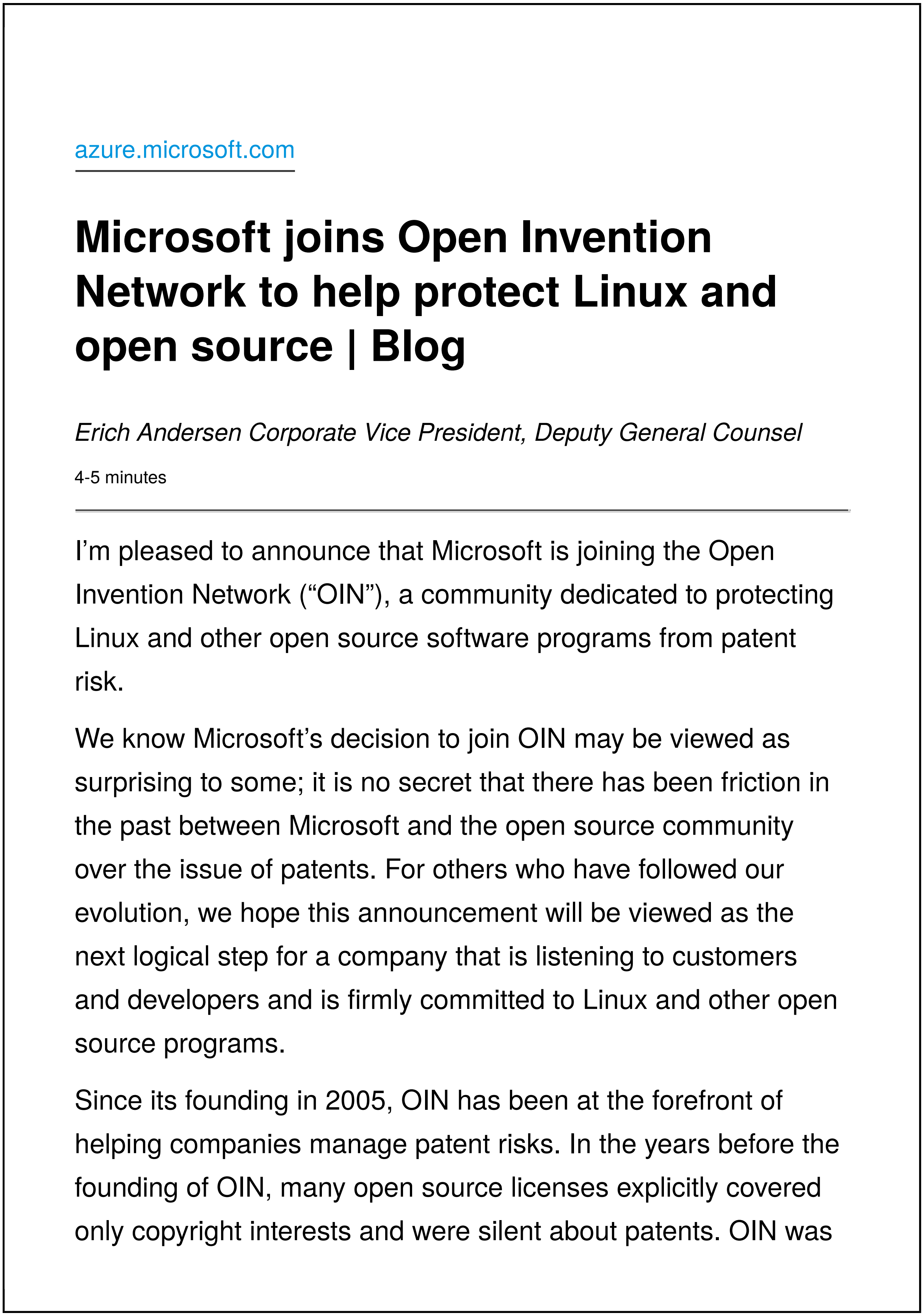


Appendix I: Language Translations

To facilitate global adoption we welcome efforts to translate the specification into multiple languages. Because OpenChain functions as an open source project translations are driven by those willing to contribute their time and expertise to perform translations under the terms of the CC-BY 4.0 license and the project's translation policy. The details of the policy and available translations can be found on the OpenChain project specification webpage.

Chapter 5

Microsoft Joins Open Invention Network to Help Protect Linux and Open Source



azure.microsoft.com

Microsoft joins Open Invention Network to help protect Linux and open source | Blog

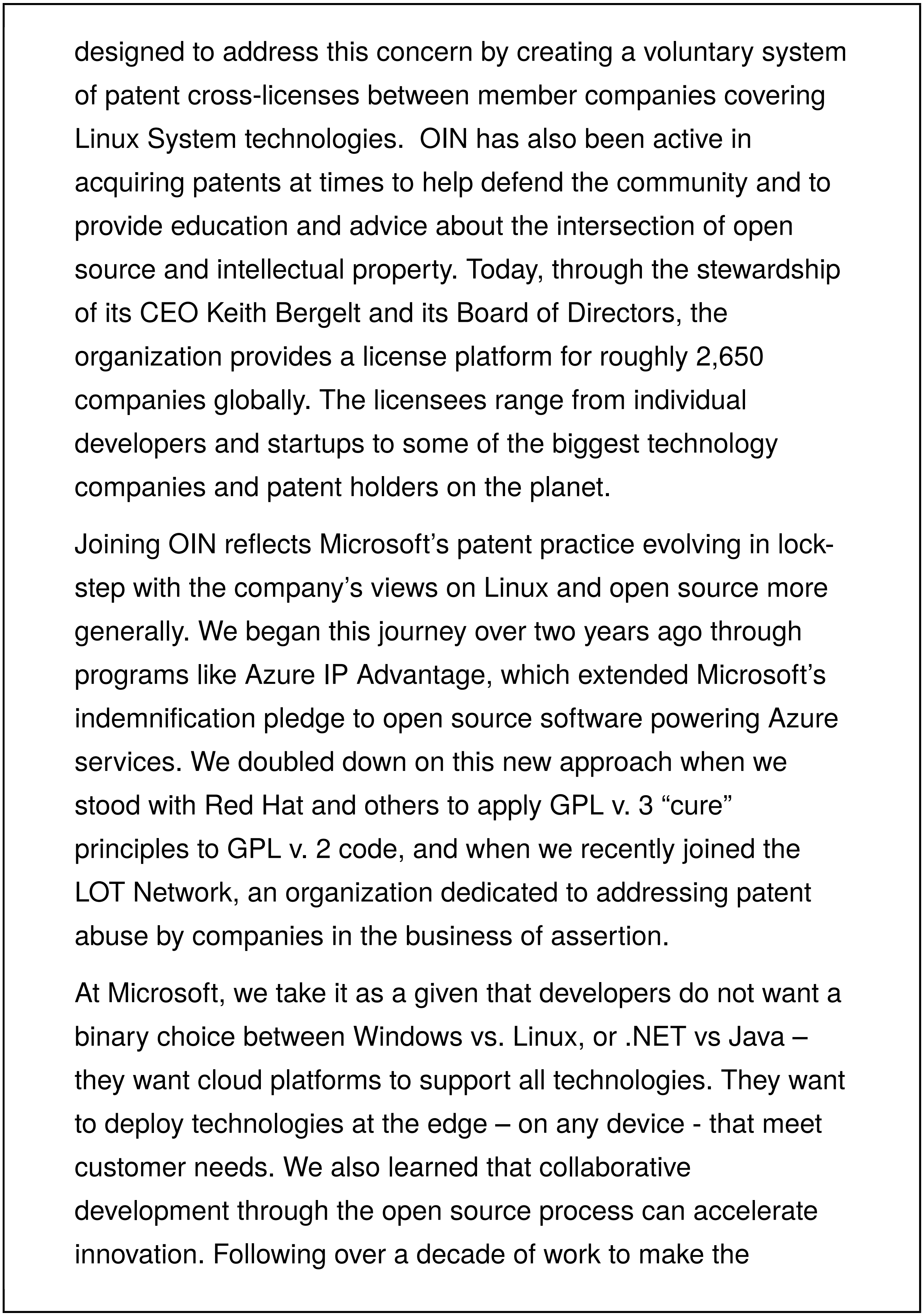
Erich Andersen Corporate Vice President, Deputy General Counsel

4-5 minutes

I'm pleased to announce that Microsoft is joining the Open Invention Network ("OIN"), a community dedicated to protecting Linux and other open source software programs from patent risk.

We know Microsoft's decision to join OIN may be viewed as surprising to some; it is no secret that there has been friction in the past between Microsoft and the open source community over the issue of patents. For others who have followed our evolution, we hope this announcement will be viewed as the next logical step for a company that is listening to customers and developers and is firmly committed to Linux and other open source programs.

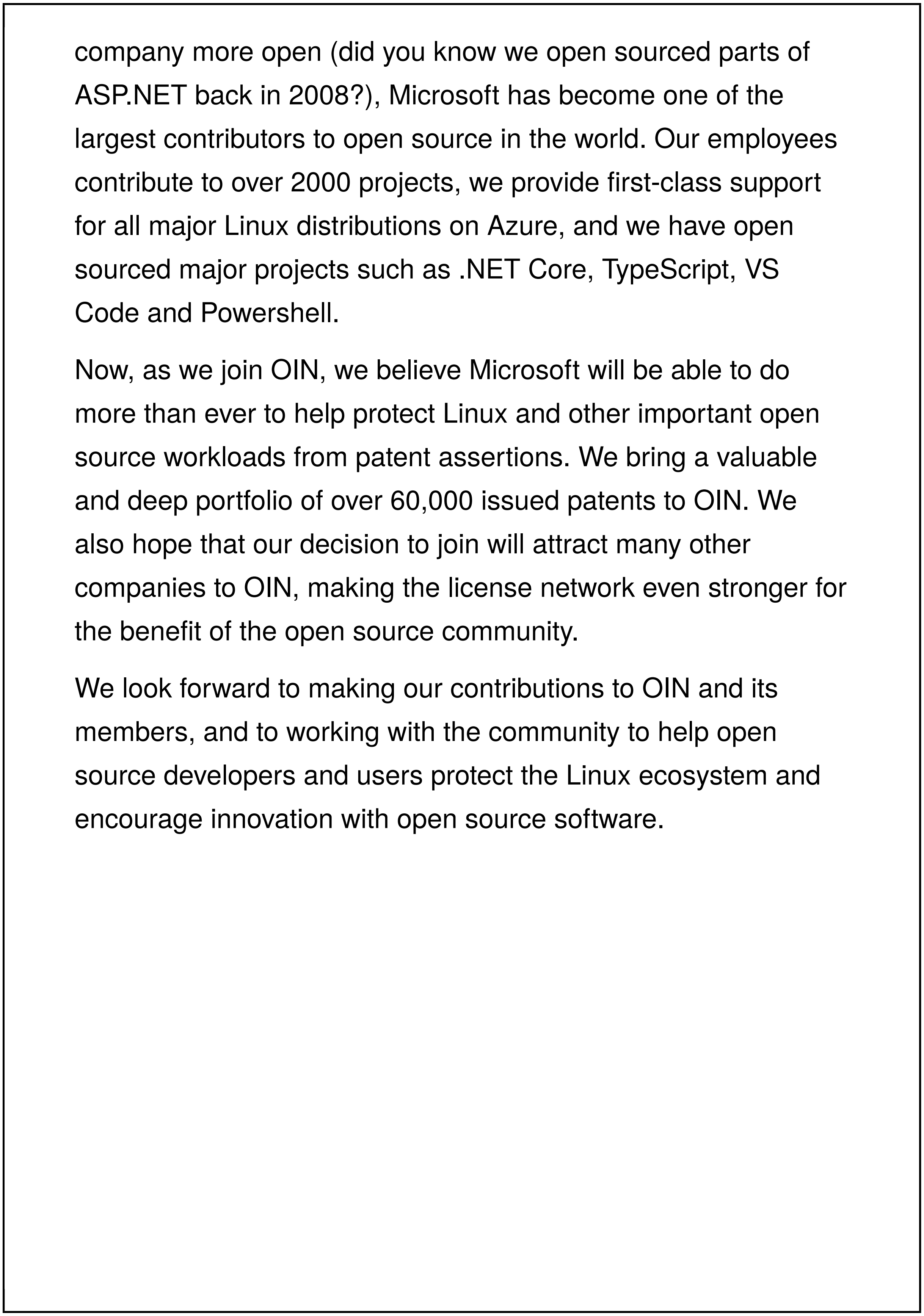
Since its founding in 2005, OIN has been at the forefront of helping companies manage patent risks. In the years before the founding of OIN, many open source licenses explicitly covered only copyright interests and were silent about patents. OIN was



designed to address this concern by creating a voluntary system of patent cross-licenses between member companies covering Linux System technologies. OIN has also been active in acquiring patents at times to help defend the community and to provide education and advice about the intersection of open source and intellectual property. Today, through the stewardship of its CEO Keith Bergelt and its Board of Directors, the organization provides a license platform for roughly 2,650 companies globally. The licensees range from individual developers and startups to some of the biggest technology companies and patent holders on the planet.

Joining OIN reflects Microsoft's patent practice evolving in lock-step with the company's views on Linux and open source more generally. We began this journey over two years ago through programs like Azure IP Advantage, which extended Microsoft's indemnification pledge to open source software powering Azure services. We doubled down on this new approach when we stood with Red Hat and others to apply GPL v. 3 "cure" principles to GPL v. 2 code, and when we recently joined the LOT Network, an organization dedicated to addressing patent abuse by companies in the business of assertion.

At Microsoft, we take it as a given that developers do not want a binary choice between Windows vs. Linux, or .NET vs Java – they want cloud platforms to support all technologies. They want to deploy technologies at the edge – on any device - that meet customer needs. We also learned that collaborative development through the open source process can accelerate innovation. Following over a decade of work to make the



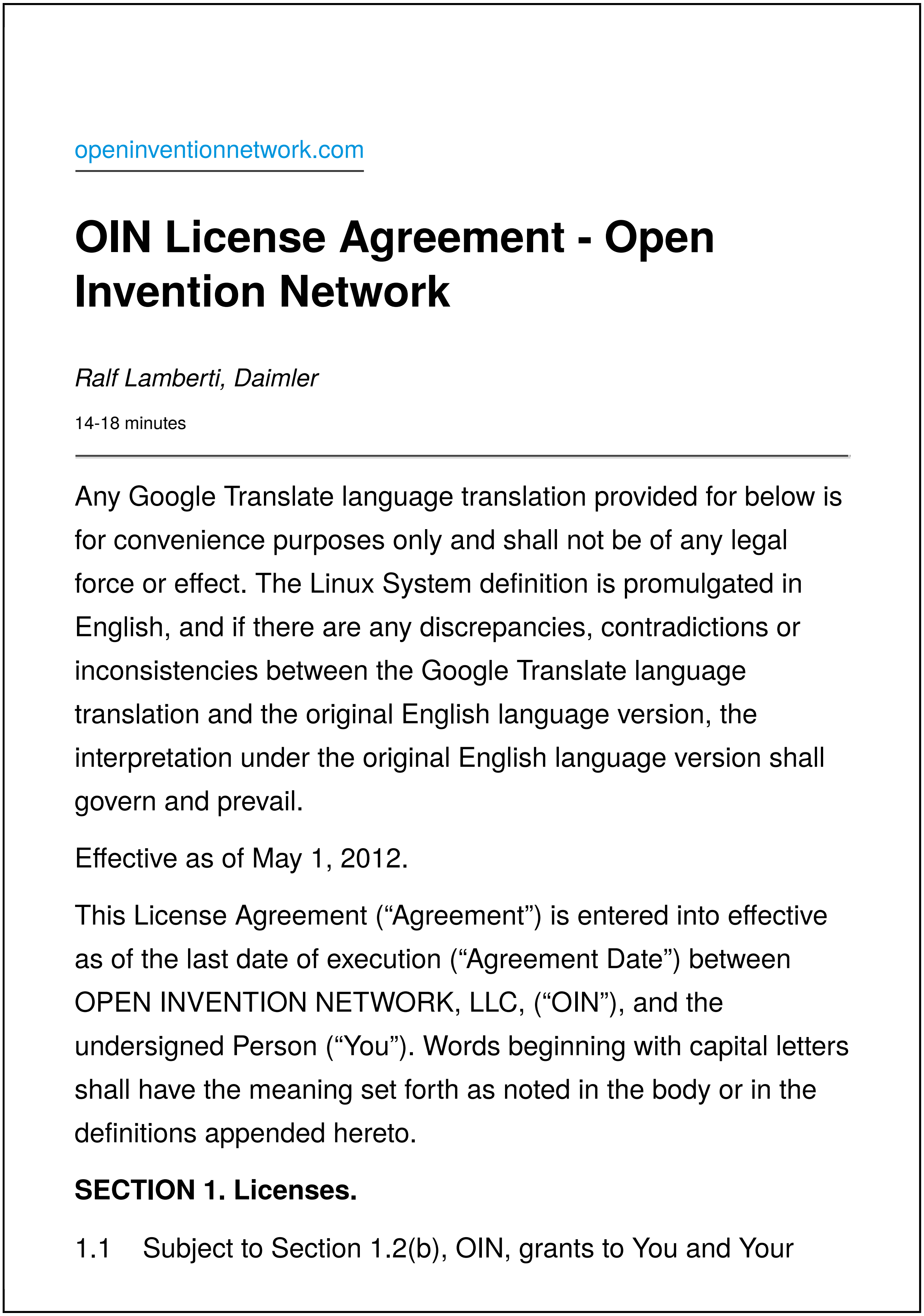
company more open (did you know we open sourced parts of ASP.NET back in 2008?), Microsoft has become one of the largest contributors to open source in the world. Our employees contribute to over 2000 projects, we provide first-class support for all major Linux distributions on Azure, and we have open sourced major projects such as .NET Core, TypeScript, VS Code and Powershell.

Now, as we join OIN, we believe Microsoft will be able to do more than ever to help protect Linux and other important open source workloads from patent assertions. We bring a valuable and deep portfolio of over 60,000 issued patents to OIN. We also hope that our decision to join will attract many other companies to OIN, making the license network even stronger for the benefit of the open source community.

We look forward to making our contributions to OIN and its members, and to working with the community to help open source developers and users protect the Linux ecosystem and encourage innovation with open source software.

Chapter 6

Open Invention Network License Agreement



openinventionnetwork.com

OIN License Agreement - Open Invention Network

Ralf Lamberti, Daimler

14-18 minutes

Any Google Translate language translation provided for below is for convenience purposes only and shall not be of any legal force or effect. The Linux System definition is promulgated in English, and if there are any discrepancies, contradictions or inconsistencies between the Google Translate language translation and the original English language version, the interpretation under the original English language version shall govern and prevail.

Effective as of May 1, 2012.

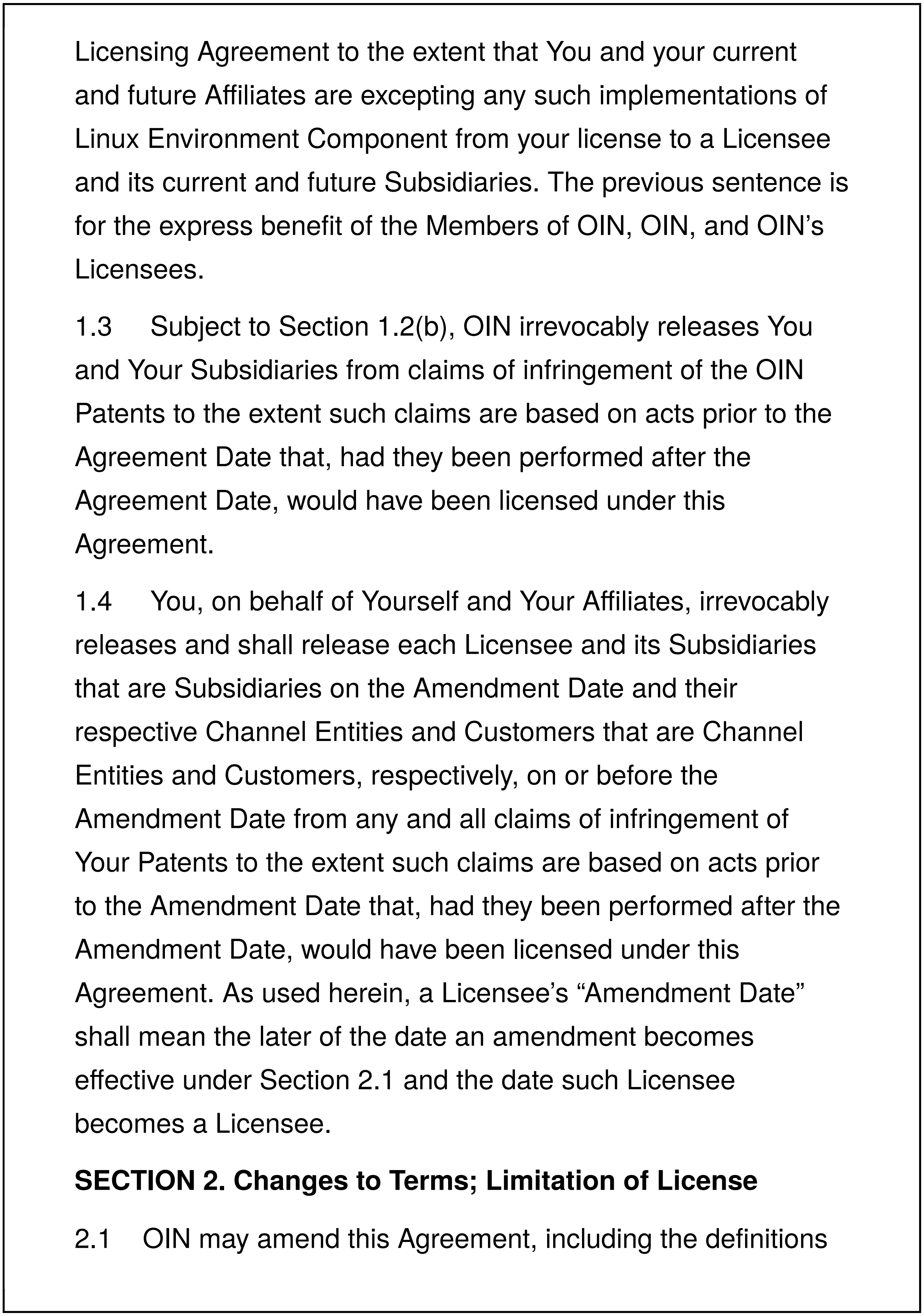
This License Agreement (“Agreement”) is entered into effective as of the last date of execution (“Agreement Date”) between OPEN INVENTION NETWORK, LLC, (“OIN”), and the undersigned Person (“You”). Words beginning with capital letters shall have the meaning set forth as noted in the body or in the definitions appended hereto.

SECTION 1. Licenses.

1.1 Subject to Section 1.2(b), OIN, grants to You and Your

Subsidiaries a royalty-free, worldwide, nonexclusive, non-transferable license under OIN Patents to make, have made, use, import, and Distribute any products or services. In addition to the foregoing and without limitation thereof, with respect only to the Linux System, the license granted herein includes the right to engage in activities that in the absence of this Agreement would constitute inducement to infringe or contributory infringement (or infringement under any other analogous legal doctrine in the applicable jurisdiction).

1.2 Subject to Section 2.2 and in consideration for the license granted in Section 1.1, You, on behalf of yourself and your Affiliates, (a) grant to each Licensee and its Subsidiaries that are Subsidiaries as of the Eligibility Date a royalty-free, worldwide, nonexclusive, non-transferable license under Your Patents for making, having made, using, importing, and Distributing any Linux System; and (b) represent and warrant that (i) You have the full right and power to grant the foregoing licenses and the release in Section 1.4 and that Your Affiliates are and will be bound by the obligations of this Agreement; and (ii) neither You nor any of Your Affiliates has a Claim pending against any Person for making, having made, using, importing, and Distributing any Linux System. Notwithstanding anything in another Company Licensing Agreement to the contrary, You and your current and future Subsidiaries do not and shall not receive, and hereby disclaim and waive, any license from a Licensee and its current and future Affiliates pursuant to a Company Licensing Agreement for implementations of Linux Environment Components as specified in such Company



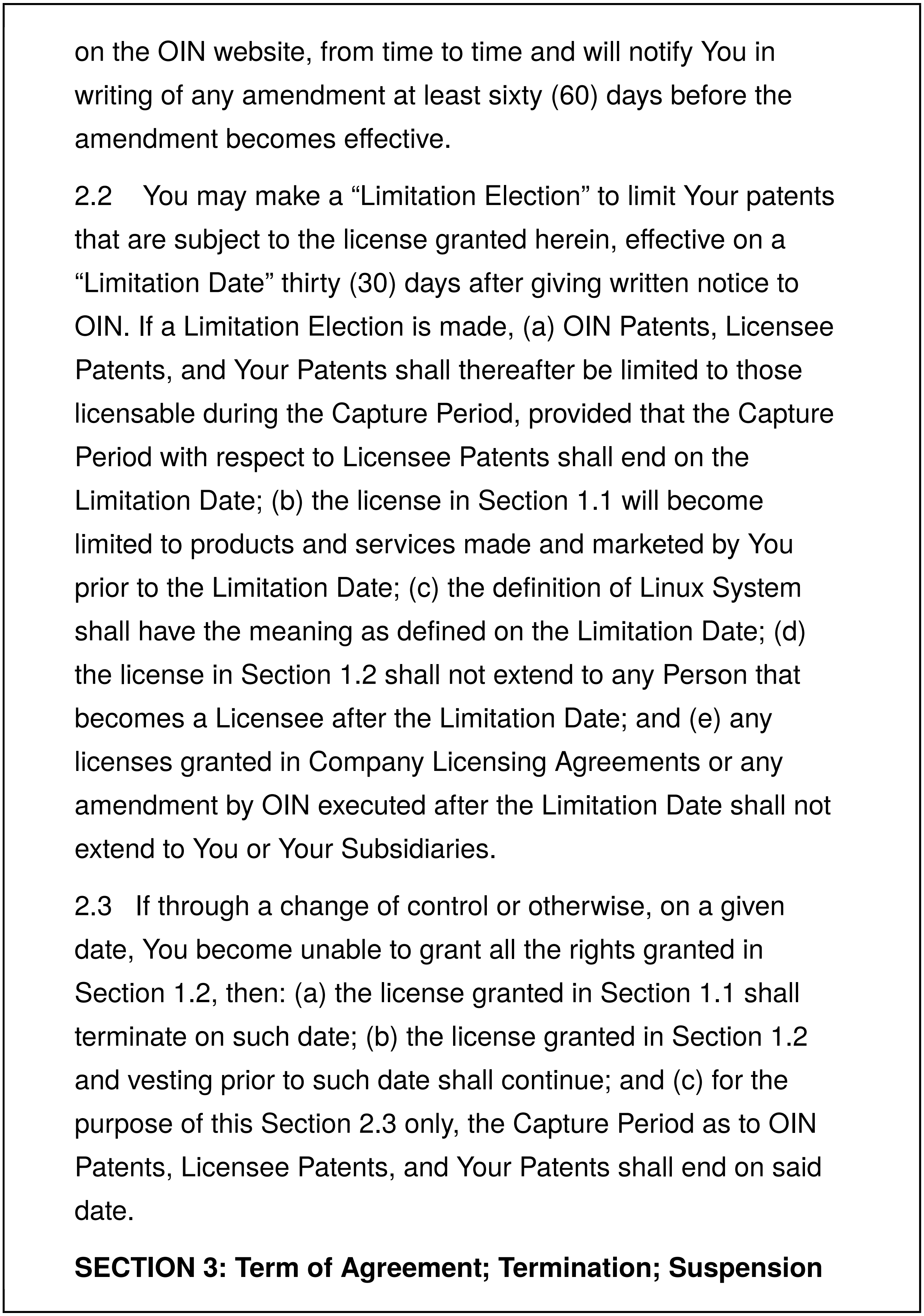
Licensing Agreement to the extent that You and your current and future Affiliates are excepting any such implementations of Linux Environment Component from your license to a Licensee and its current and future Subsidiaries. The previous sentence is for the express benefit of the Members of OIN, OIN, and OIN's Licensees.

1.3 Subject to Section 1.2(b), OIN irrevocably releases You and Your Subsidiaries from claims of infringement of the OIN Patents to the extent such claims are based on acts prior to the Agreement Date that, had they been performed after the Agreement Date, would have been licensed under this Agreement.

1.4 You, on behalf of Yourself and Your Affiliates, irrevocably releases and shall release each Licensee and its Subsidiaries that are Subsidiaries on the Amendment Date and their respective Channel Entities and Customers that are Channel Entities and Customers, respectively, on or before the Amendment Date from any and all claims of infringement of Your Patents to the extent such claims are based on acts prior to the Amendment Date that, had they been performed after the Amendment Date, would have been licensed under this Agreement. As used herein, a Licensee's "Amendment Date" shall mean the later of the date an amendment becomes effective under Section 2.1 and the date such Licensee becomes a Licensee.

SECTION 2. Changes to Terms; Limitation of License

2.1 OIN may amend this Agreement, including the definitions



on the OIN website, from time to time and will notify You in writing of any amendment at least sixty (60) days before the amendment becomes effective.

2.2 You may make a “Limitation Election” to limit Your patents that are subject to the license granted herein, effective on a “Limitation Date” thirty (30) days after giving written notice to OIN. If a Limitation Election is made, (a) OIN Patents, Licensee Patents, and Your Patents shall thereafter be limited to those licensable during the Capture Period, provided that the Capture Period with respect to Licensee Patents shall end on the Limitation Date; (b) the license in Section 1.1 will become limited to products and services made and marketed by You prior to the Limitation Date; (c) the definition of Linux System shall have the meaning as defined on the Limitation Date; (d) the license in Section 1.2 shall not extend to any Person that becomes a Licensee after the Limitation Date; and (e) any licenses granted in Company Licensing Agreements or any amendment by OIN executed after the Limitation Date shall not extend to You or Your Subsidiaries.

2.3 If through a change of control or otherwise, on a given date, You become unable to grant all the rights granted in Section 1.2, then: (a) the license granted in Section 1.1 shall terminate on such date; (b) the license granted in Section 1.2 and vesting prior to such date shall continue; and (c) for the purpose of this Section 2.3 only, the Capture Period as to OIN Patents, Licensee Patents, and Your Patents shall end on said date.

SECTION 3: Term of Agreement; Termination; Suspension

3.1 The term of this Agreement shall be from the Agreement Date until the last to expire of the OIN Patents or Your Patents, unless earlier terminated.

3.2 If a Subsidiary of You ceases to be a Subsidiary on a given date, the license granted in Section 1.1 to such Subsidiary shall terminate on such date. If an Affiliate of You ceases to be an Affiliate on a given date, the license granted in Section 1.2 and vesting prior to such date by such Affiliate shall continue.

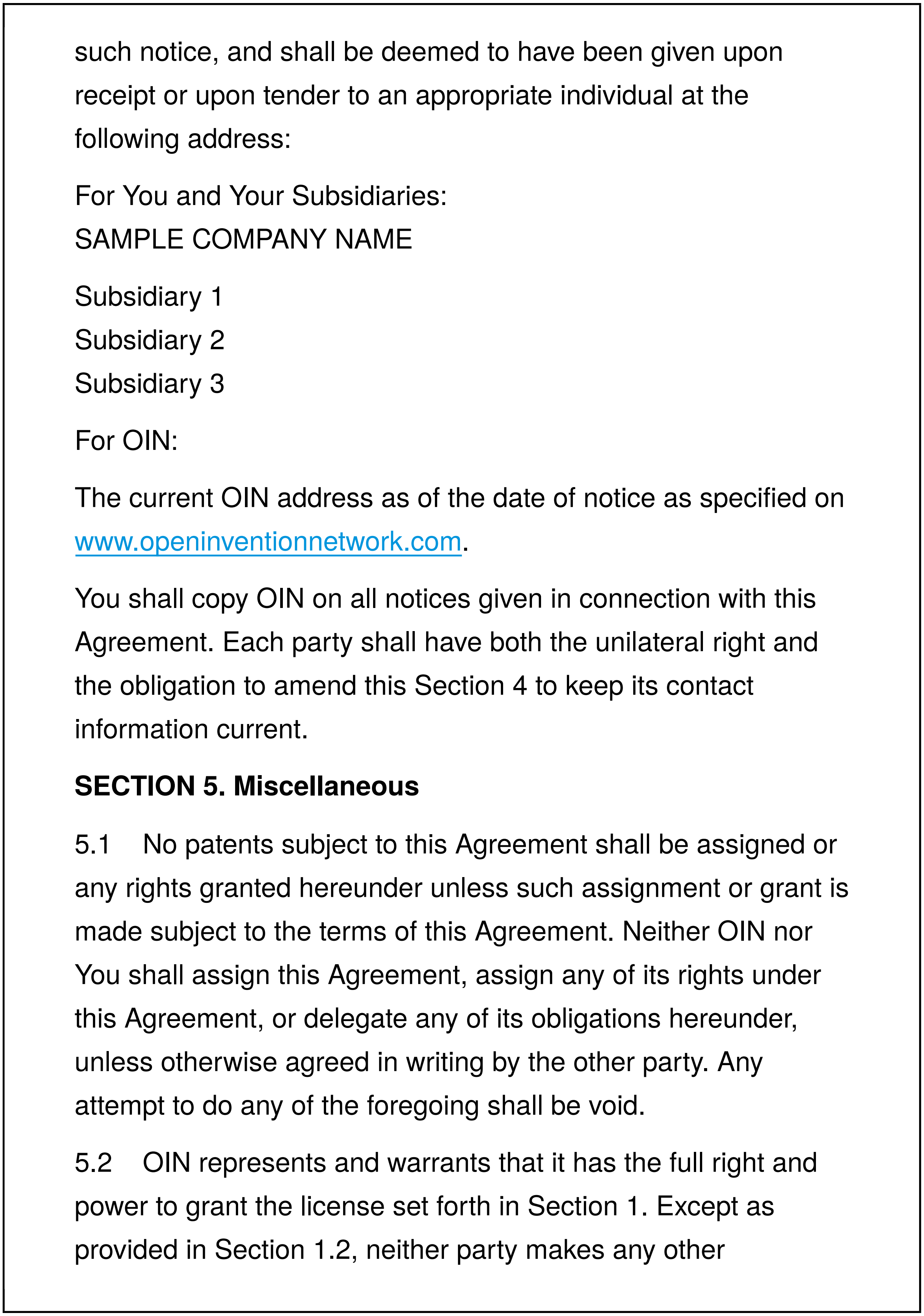
3.3 If a Licensee or its Affiliate files one or more Claims against You or Your Subsidiaries based on products that perform substantially the same function as the Linux System, and are Distributed by You or Your Subsidiaries, then You may suspend the license granted under Section 1.2 to such Licensee and its Subsidiaries on written notice to such Licensee. Such suspension shall be effective unless and until such Claim is dismissed.

3.4 The license in Section 1.1 shall terminate effective on the day You or Your Subsidiary files one or more Claims against any Licensee, whose license has not been suspended by You under Section 3.3, for making, having made, using, importing, or Distributing any Linux System.

3.5 No termination or suspension of the licenses granted hereunder shall relieve either party of any obligation accrued hereunder prior to such termination.

SECTION 4: Notice

Notices and other communications in connection with this Agreement shall be in writing and signed by the party giving



such notice, and shall be deemed to have been given upon receipt or upon tender to an appropriate individual at the following address:

For You and Your Subsidiaries:
SAMPLE COMPANY NAME

Subsidiary 1
Subsidiary 2
Subsidiary 3

For OIN:

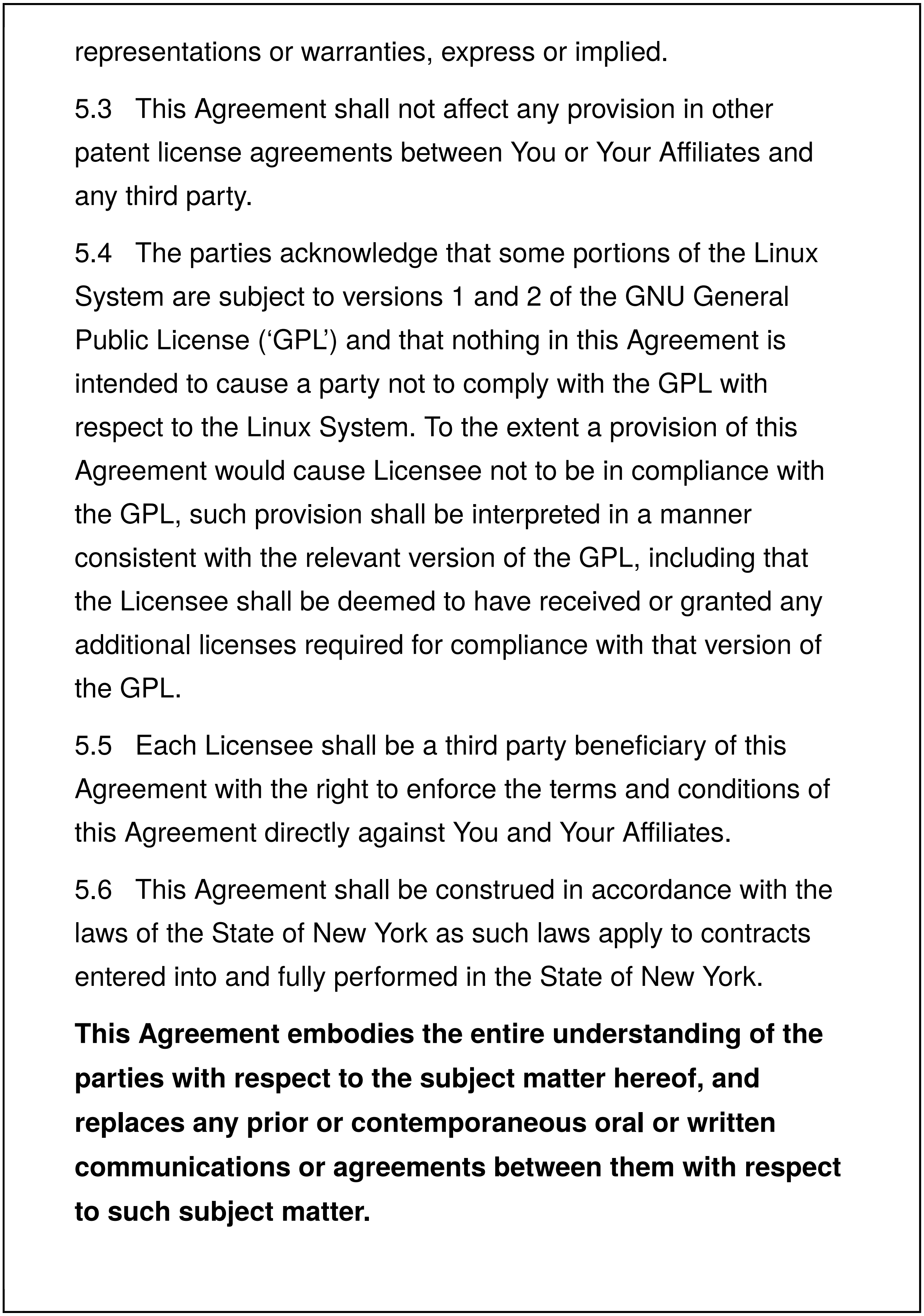
The current OIN address as of the date of notice as specified on www.openinventionnetwork.com.

You shall copy OIN on all notices given in connection with this Agreement. Each party shall have both the unilateral right and the obligation to amend this Section 4 to keep its contact information current.

SECTION 5. Miscellaneous

5.1 No patents subject to this Agreement shall be assigned or any rights granted hereunder unless such assignment or grant is made subject to the terms of this Agreement. Neither OIN nor You shall assign this Agreement, assign any of its rights under this Agreement, or delegate any of its obligations hereunder, unless otherwise agreed in writing by the other party. Any attempt to do any of the foregoing shall be void.

5.2 OIN represents and warrants that it has the full right and power to grant the license set forth in Section 1. Except as provided in Section 1.2, neither party makes any other



representations or warranties, express or implied.

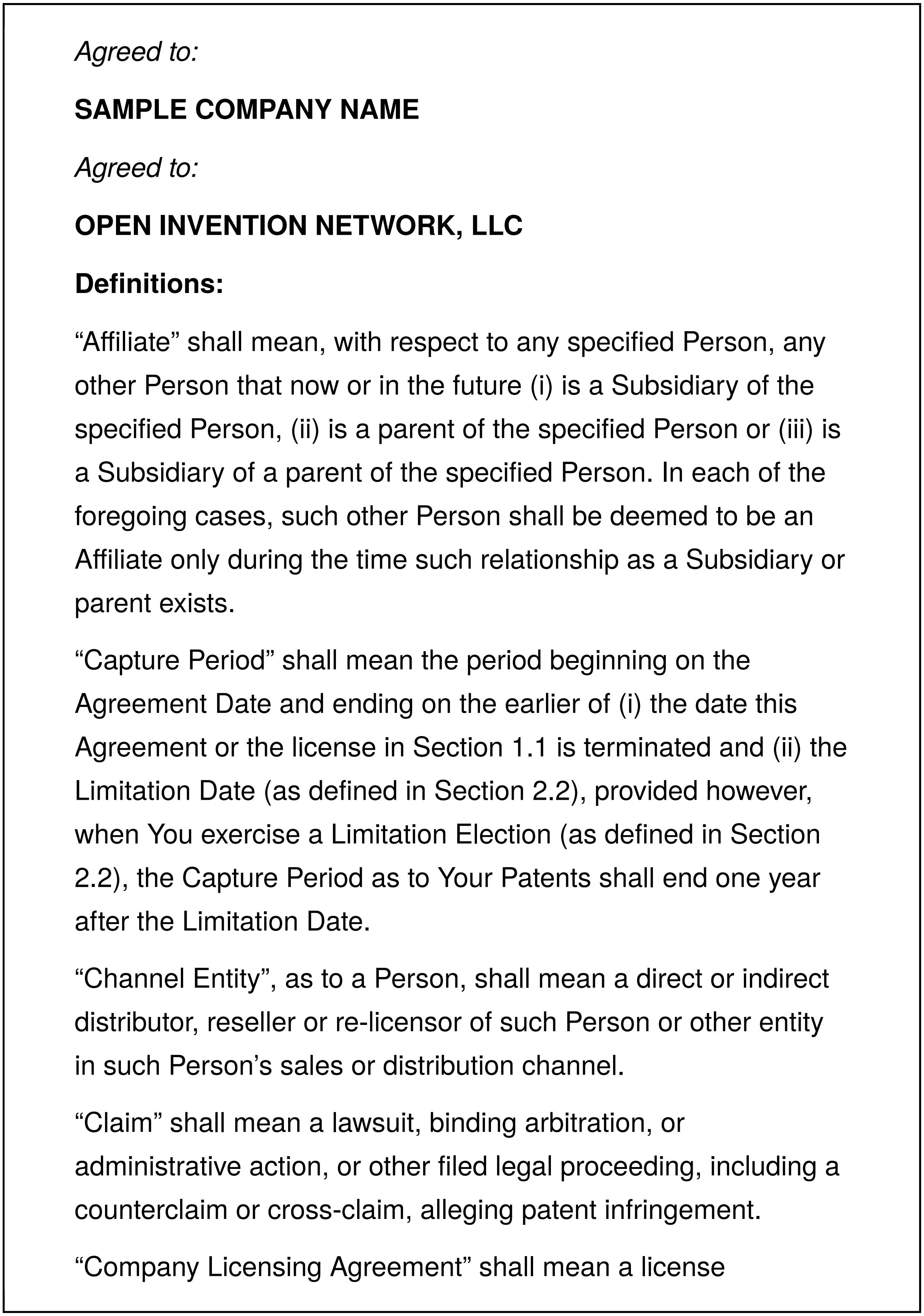
5.3 This Agreement shall not affect any provision in other patent license agreements between You or Your Affiliates and any third party.

5.4 The parties acknowledge that some portions of the Linux System are subject to versions 1 and 2 of the GNU General Public License (‘GPL’) and that nothing in this Agreement is intended to cause a party not to comply with the GPL with respect to the Linux System. To the extent a provision of this Agreement would cause Licensee not to be in compliance with the GPL, such provision shall be interpreted in a manner consistent with the relevant version of the GPL, including that the Licensee shall be deemed to have received or granted any additional licenses required for compliance with that version of the GPL.

5.5 Each Licensee shall be a third party beneficiary of this Agreement with the right to enforce the terms and conditions of this Agreement directly against You and Your Affiliates.

5.6 This Agreement shall be construed in accordance with the laws of the State of New York as such laws apply to contracts entered into and fully performed in the State of New York.

This Agreement embodies the entire understanding of the parties with respect to the subject matter hereof, and replaces any prior or contemporaneous oral or written communications or agreements between them with respect to such subject matter.



Agreed to:

SAMPLE COMPANY NAME

Agreed to:

OPEN INVENTION NETWORK, LLC

Definitions:

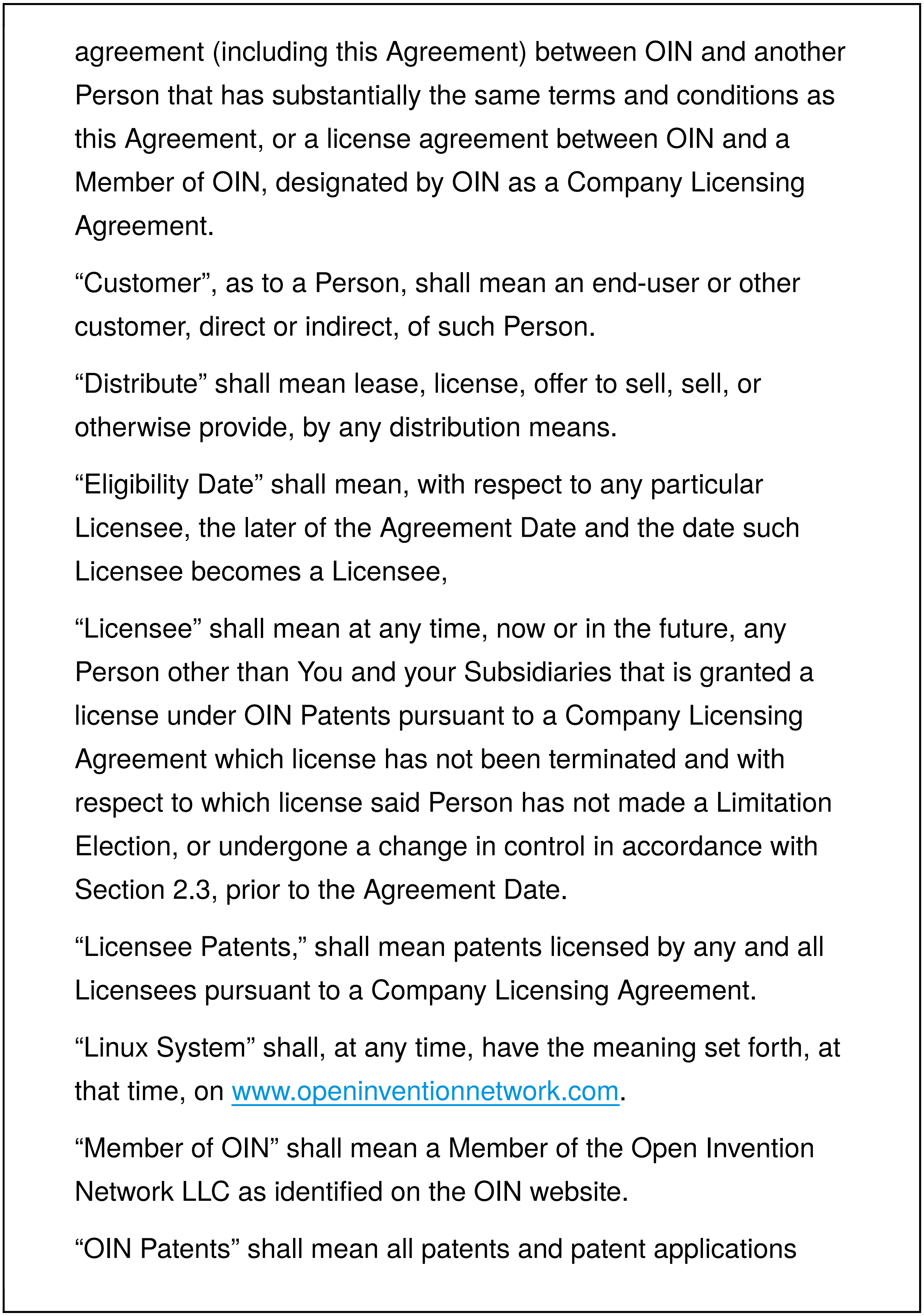
“Affiliate” shall mean, with respect to any specified Person, any other Person that now or in the future (i) is a Subsidiary of the specified Person, (ii) is a parent of the specified Person or (iii) is a Subsidiary of a parent of the specified Person. In each of the foregoing cases, such other Person shall be deemed to be an Affiliate only during the time such relationship as a Subsidiary or parent exists.

“Capture Period” shall mean the period beginning on the Agreement Date and ending on the earlier of (i) the date this Agreement or the license in Section 1.1 is terminated and (ii) the Limitation Date (as defined in Section 2.2), provided however, when You exercise a Limitation Election (as defined in Section 2.2), the Capture Period as to Your Patents shall end one year after the Limitation Date.

“Channel Entity”, as to a Person, shall mean a direct or indirect distributor, reseller or re-licensor of such Person or other entity in such Person’s sales or distribution channel.

“Claim” shall mean a lawsuit, binding arbitration, or administrative action, or other filed legal proceeding, including a counterclaim or cross-claim, alleging patent infringement.

“Company Licensing Agreement” shall mean a license



agreement (including this Agreement) between OIN and another Person that has substantially the same terms and conditions as this Agreement, or a license agreement between OIN and a Member of OIN, designated by OIN as a Company Licensing Agreement.

“Customer”, as to a Person, shall mean an end-user or other customer, direct or indirect, of such Person.

“Distribute” shall mean lease, license, offer to sell, sell, or otherwise provide, by any distribution means.

“Eligibility Date” shall mean, with respect to any particular Licensee, the later of the Agreement Date and the date such Licensee becomes a Licensee,

“Licensee” shall mean at any time, now or in the future, any Person other than You and your Subsidiaries that is granted a license under OIN Patents pursuant to a Company Licensing Agreement which license has not been terminated and with respect to which license said Person has not made a Limitation Election, or undergone a change in control in accordance with Section 2.3, prior to the Agreement Date.

“Licensee Patents,” shall mean patents licensed by any and all Licensees pursuant to a Company Licensing Agreement.

“Linux System” shall, at any time, have the meaning set forth, at that time, on www.openinventionnetwork.com.

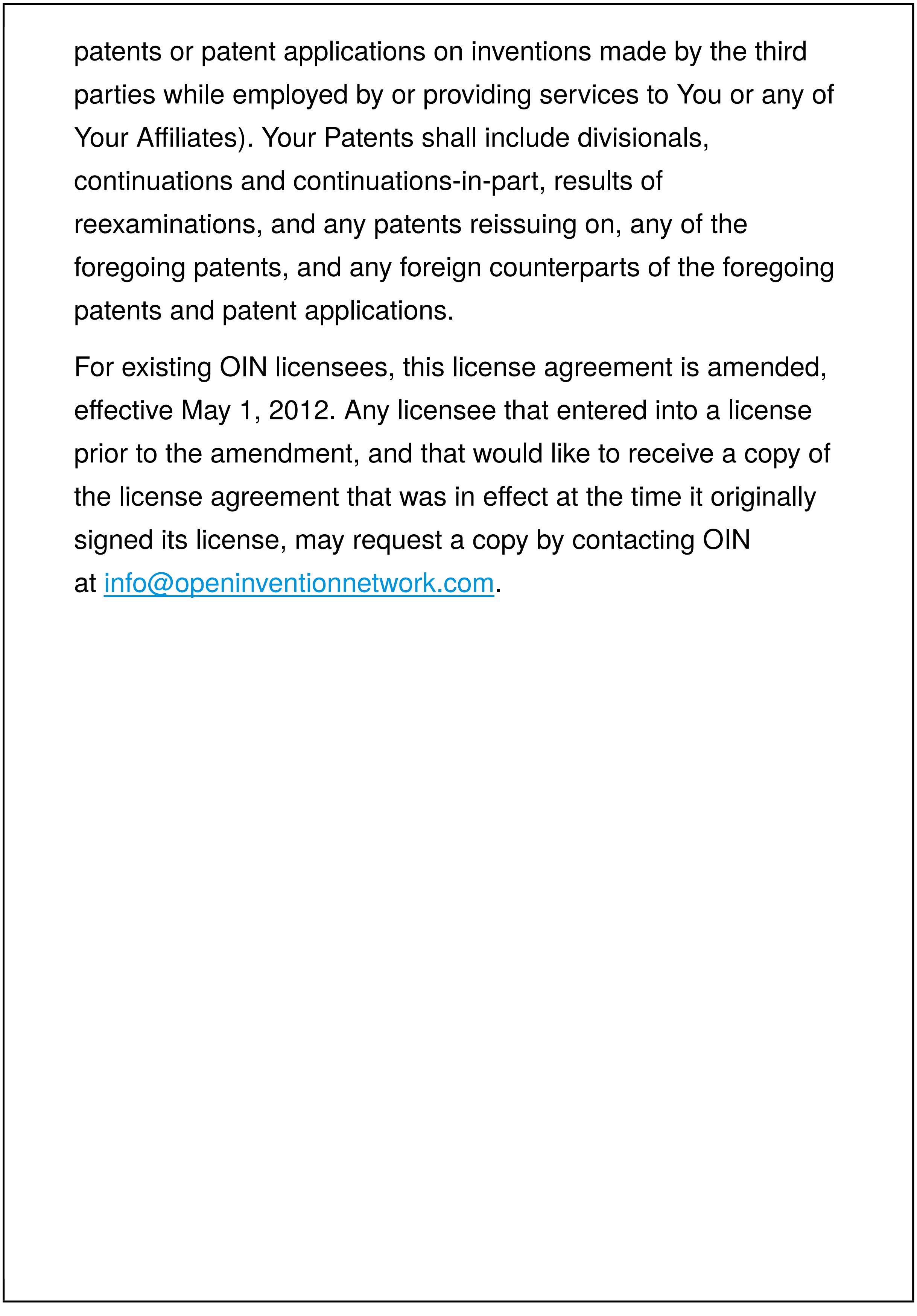
“Member of OIN” shall mean a Member of the Open Invention Network LLC as identified on the OIN website.

“OIN Patents” shall mean all patents and patent applications including utility models and typeface design patents and registrations, under which OIN has at any time during the Capture Period, the right to grant licenses to You or Your Subsidiaries of or within the scope granted herein without such grant or the exercise of rights thereunder resulting in the payment of royalties or other consideration by OIN to unaffiliated third parties. OIN Patents shall include divisionals, continuations and continuations-in-part, results of reexaminations, any foreign counterparts of the foregoing patents and patent applications and any patents reissuing on any of the foregoing patents.

“Person” includes any individual, corporation, association, partnership (general or limited), joint venture, trust, estate, limited liability company or other legal entity or organization.

“Subsidiary” shall mean, with respect to any specified Person, any other Person of which more than 50% of the total voting power is owned or controlled, directly or indirectly, now or in the future, by the specified Person, but such other Person shall be deemed to be a Subsidiary only during the time such ownership or control exists.

“Your Patents” shall mean all patents and patent applications including utility models and typeface design patents and registrations (but not including any other design patents or registrations), under which You or any of Your Affiliates has at any time during the Capture Period, the right to grant rights of or within the scope granted herein without such grant or the exercise of rights thereunder resulting in the payment of royalties or other consideration by You or Your Affiliates to unaffiliated third parties (other than payments to third parties for

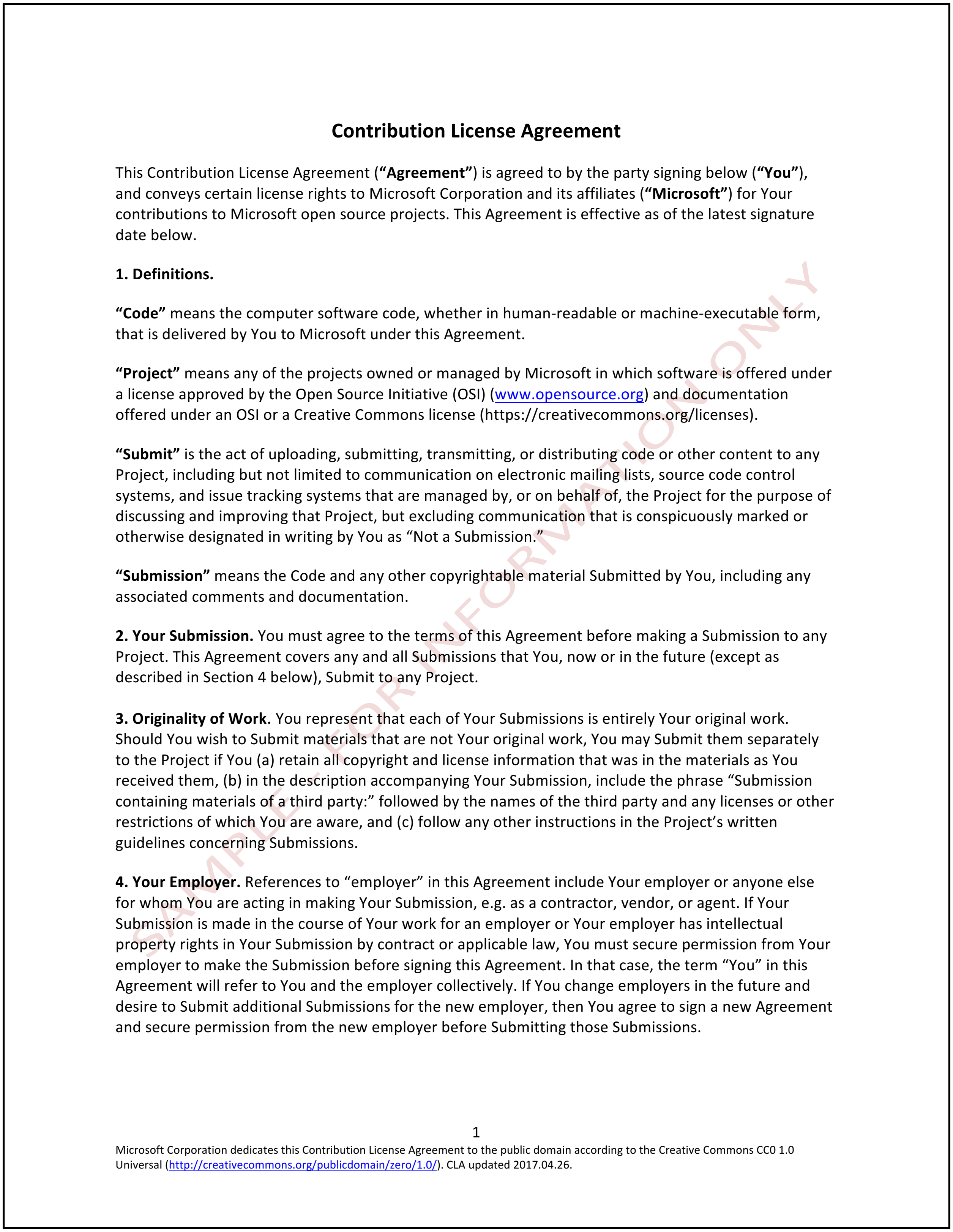


patents or patent applications on inventions made by the third parties while employed by or providing services to You or any of Your Affiliates). Your Patents shall include divisionals, continuations and continuations-in-part, results of reexaminations, and any patents reissuing on, any of the foregoing patents, and any foreign counterparts of the foregoing patents and patent applications.

For existing OIN licensees, this license agreement is amended, effective May 1, 2012. Any licensee that entered into a license prior to the amendment, and that would like to receive a copy of the license agreement that was in effect at the time it originally signed its license, may request a copy by contacting OIN at info@openinventionnetwork.com.

Chapter 7

Microsoft Open Source License Agreement



Contribution License Agreement

This Contribution License Agreement (**“Agreement”**) is agreed to by the party signing below (**“You”**), and conveys certain license rights to Microsoft Corporation and its affiliates (**“Microsoft”**) for Your contributions to Microsoft open source projects. This Agreement is effective as of the latest signature date below.

1. Definitions.

“Code” means the computer software code, whether in human-readable or machine-executable form, that is delivered by You to Microsoft under this Agreement.

“Project” means any of the projects owned or managed by Microsoft in which software is offered under a license approved by the Open Source Initiative (OSI) (www.opensource.org) and documentation offered under an OSI or a Creative Commons license (https://creativecommons.org/licenses).

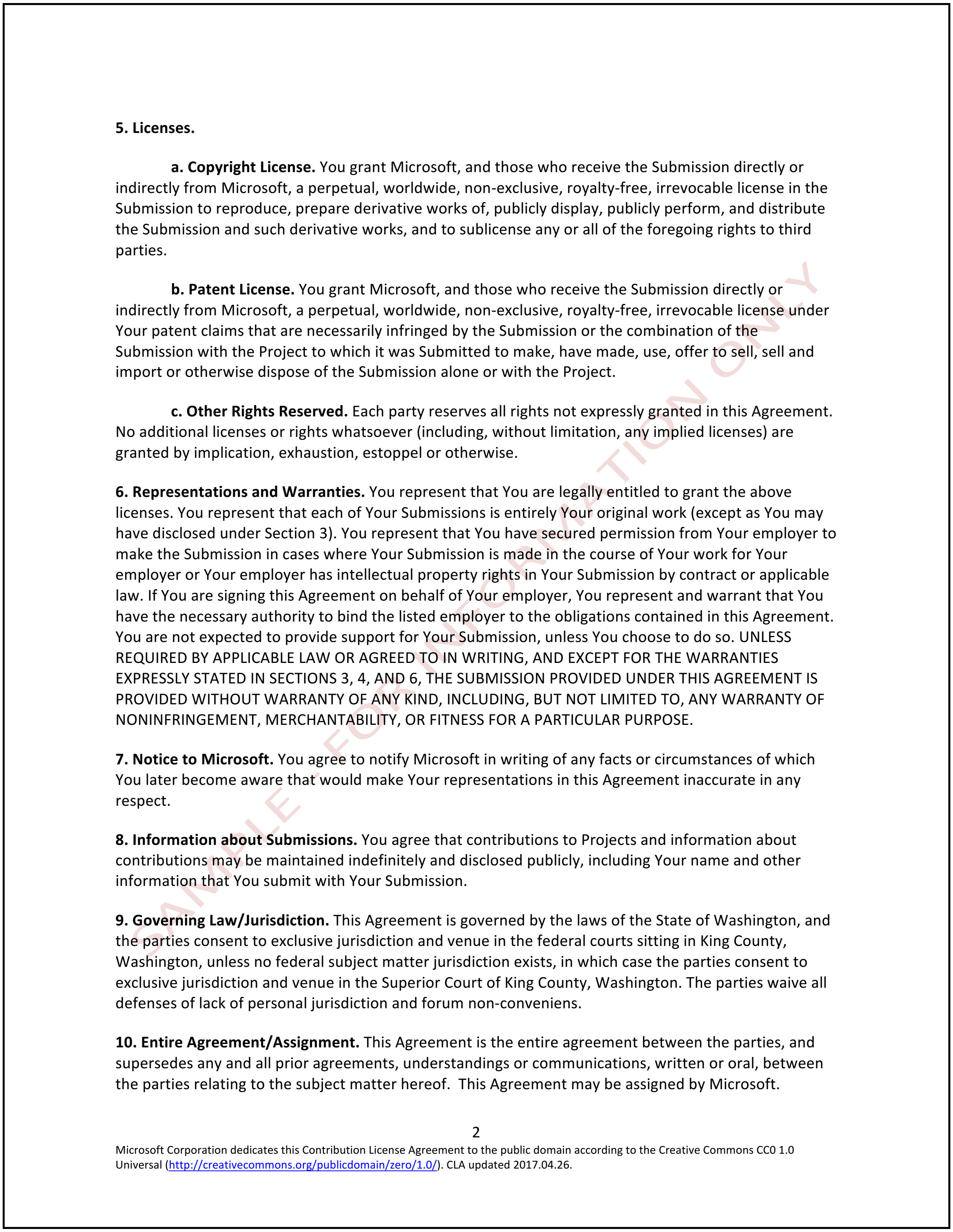
“Submit” is the act of uploading, submitting, transmitting, or distributing code or other content to any Project, including but not limited to communication on electronic mailing lists, source code control systems, and issue tracking systems that are managed by, or on behalf of, the Project for the purpose of discussing and improving that Project, but excluding communication that is conspicuously marked or otherwise designated in writing by You as “Not a Submission.”

“Submission” means the Code and any other copyrightable material Submitted by You, including any associated comments and documentation.

2. Your Submission. You must agree to the terms of this Agreement before making a Submission to any Project. This Agreement covers any and all Submissions that You, now or in the future (except as described in Section 4 below), Submit to any Project.

3. Originality of Work. You represent that each of Your Submissions is entirely Your original work. Should You wish to Submit materials that are not Your original work, You may Submit them separately to the Project if You (a) retain all copyright and license information that was in the materials as You received them, (b) in the description accompanying Your Submission, include the phrase “Submission containing materials of a third party:” followed by the names of the third party and any licenses or other restrictions of which You are aware, and (c) follow any other instructions in the Project’s written guidelines concerning Submissions.

4. Your Employer. References to “employer” in this Agreement include Your employer or anyone else for whom You are acting in making Your Submission, e.g. as a contractor, vendor, or agent. If Your Submission is made in the course of Your work for an employer or Your employer has intellectual property rights in Your Submission by contract or applicable law, You must secure permission from Your employer to make the Submission before signing this Agreement. In that case, the term “You” in this Agreement will refer to You and the employer collectively. If You change employers in the future and desire to Submit additional Submissions for the new employer, then You agree to sign a new Agreement and secure permission from the new employer before Submitting those Submissions.



5. Licenses.

a. Copyright License. You grant Microsoft, and those who receive the Submission directly or indirectly from Microsoft, a perpetual, worldwide, non-exclusive, royalty-free, irrevocable license in the Submission to reproduce, prepare derivative works of, publicly display, publicly perform, and distribute the Submission and such derivative works, and to sublicense any or all of the foregoing rights to third parties.

b. Patent License. You grant Microsoft, and those who receive the Submission directly or indirectly from Microsoft, a perpetual, worldwide, non-exclusive, royalty-free, irrevocable license under Your patent claims that are necessarily infringed by the Submission or the combination of the Submission with the Project to which it was Submitted to make, have made, use, offer to sell, sell and import or otherwise dispose of the Submission alone or with the Project.

c. Other Rights Reserved. Each party reserves all rights not expressly granted in this Agreement. No additional licenses or rights whatsoever (including, without limitation, any implied licenses) are granted by implication, exhaustion, estoppel or otherwise.

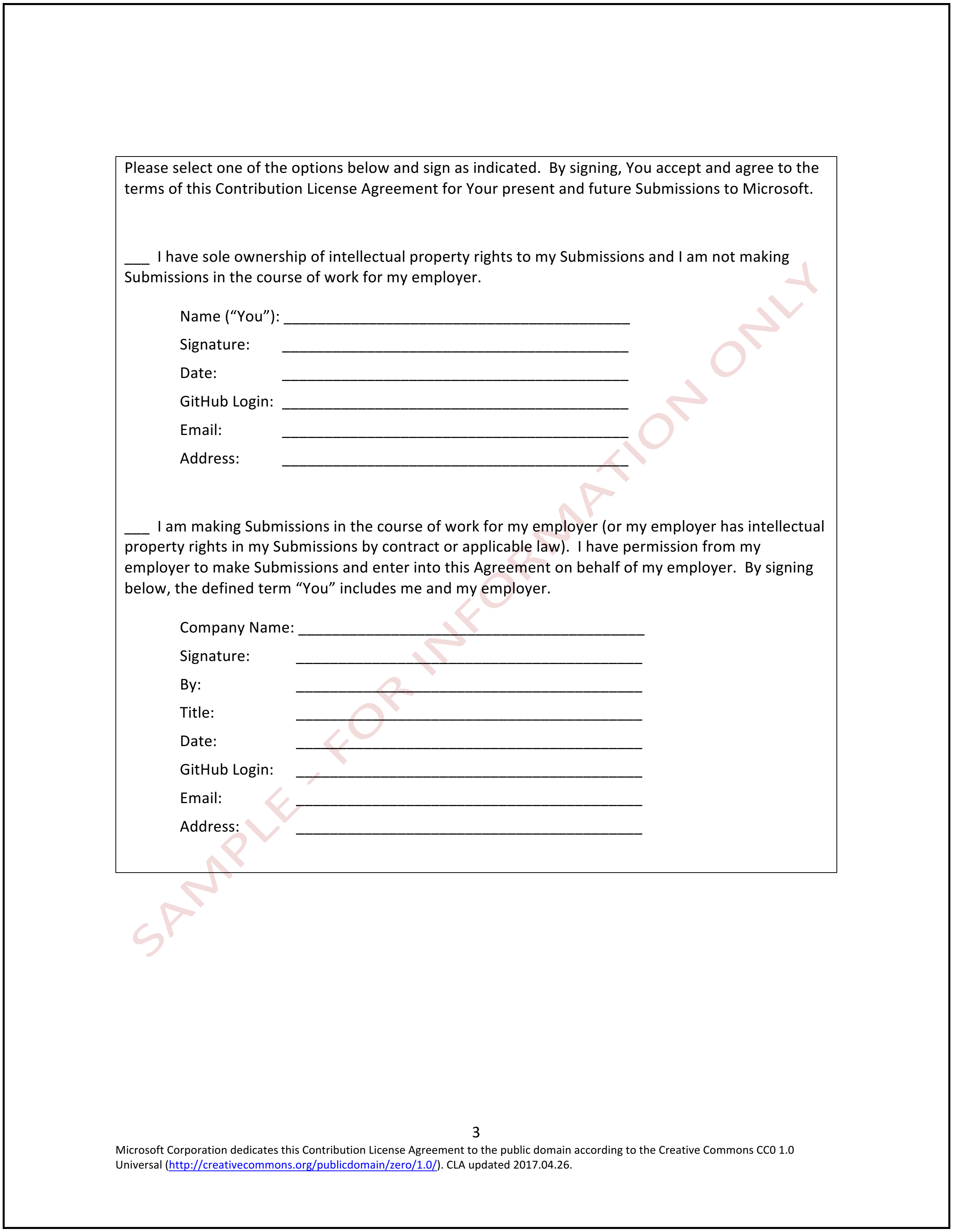
6. Representations and Warranties. You represent that You are legally entitled to grant the above licenses. You represent that each of Your Submissions is entirely Your original work (except as You may have disclosed under Section 3). You represent that You have secured permission from Your employer to make the Submission in cases where Your Submission is made in the course of Your work for Your employer or Your employer has intellectual property rights in Your Submission by contract or applicable law. If You are signing this Agreement on behalf of Your employer, You represent and warrant that You have the necessary authority to bind the listed employer to the obligations contained in this Agreement. You are not expected to provide support for Your Submission, unless You choose to do so. UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING, AND EXCEPT FOR THE WARRANTIES EXPRESSLY STATED IN SECTIONS 3, 4, AND 6, THE SUBMISSION PROVIDED UNDER THIS AGREEMENT IS PROVIDED WITHOUT WARRANTY OF ANY KIND, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTY OF NONINFRINGEMENT, MERCHANTABILITY, OR FITNESS FOR A PARTICULAR PURPOSE.

7. Notice to Microsoft. You agree to notify Microsoft in writing of any facts or circumstances of which You later become aware that would make Your representations in this Agreement inaccurate in any respect.

8. Information about Submissions. You agree that contributions to Projects and information about contributions may be maintained indefinitely and disclosed publicly, including Your name and other information that You submit with Your Submission.

9. Governing Law/Jurisdiction. This Agreement is governed by the laws of the State of Washington, and the parties consent to exclusive jurisdiction and venue in the federal courts sitting in King County, Washington, unless no federal subject matter jurisdiction exists, in which case the parties consent to exclusive jurisdiction and venue in the Superior Court of King County, Washington. The parties waive all defenses of lack of personal jurisdiction and forum non-conveniens.

10. Entire Agreement/Assignment. This Agreement is the entire agreement between the parties, and supersedes any and all prior agreements, understandings or communications, written or oral, between the parties relating to the subject matter hereof. This Agreement may be assigned by Microsoft.



Please select one of the options below and sign as indicated. By signing, You accept and agree to the terms of this Contribution License Agreement for Your present and future Submissions to Microsoft.

___ I have sole ownership of intellectual property rights to my Submissions and I am not making Submissions in the course of work for my employer.

Name ("You"): ____

Signature: ____

Date: ____

GitHub Login: ____

Email: ____

Address: ____

___ I am making Submissions in the course of work for my employer (or my employer has intellectual property rights in my Submissions by contract or applicable law). I have permission from my employer to make Submissions and enter into this Agreement on behalf of my employer. By signing below, the defined term "You" includes me and my employer.

Company Name: ____

Signature: ____

By: ____

Title: ____

Date: ____

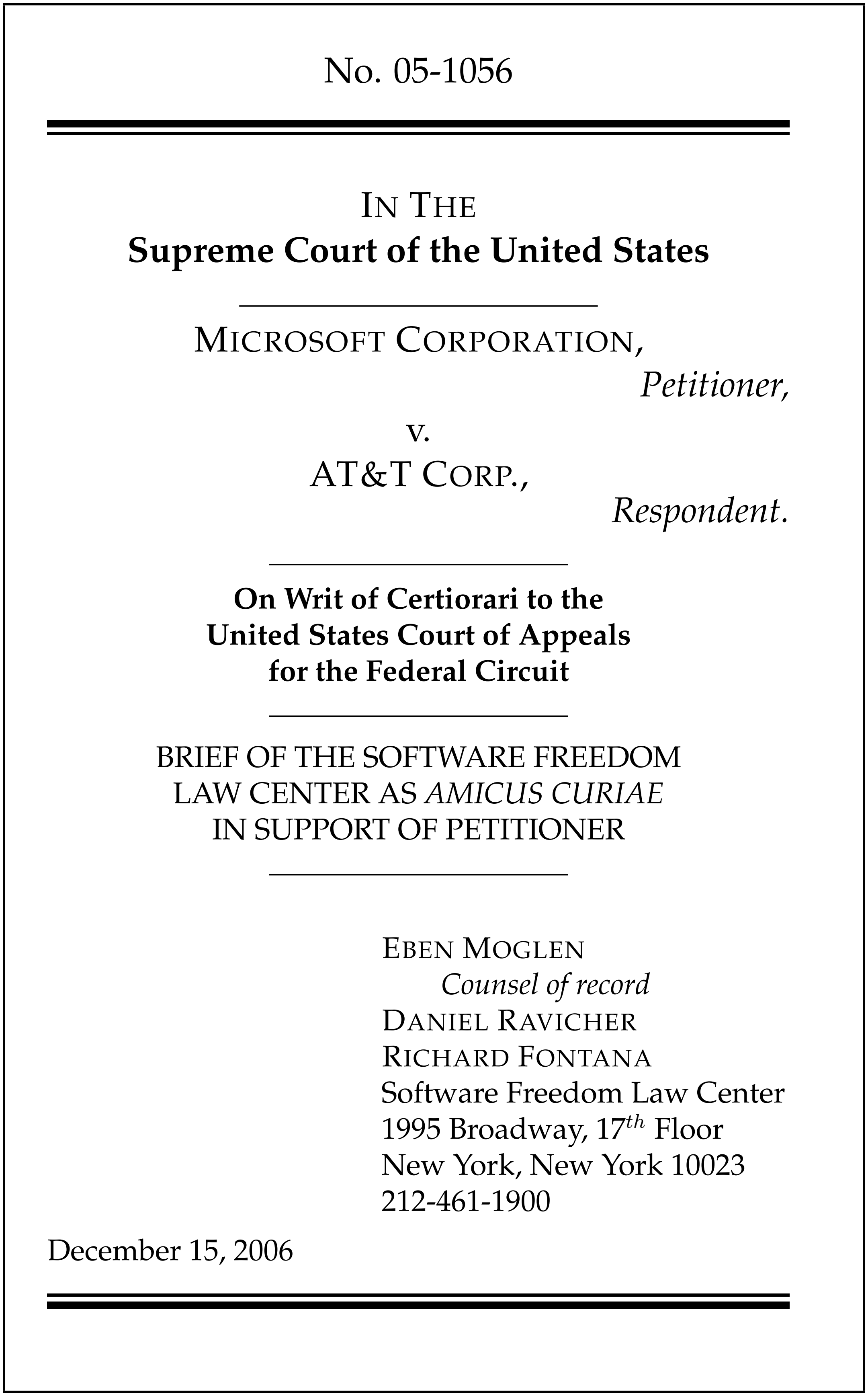
GitHub Login: ____

Email: ____

Address: ____

Chapter 8

Microsoft v. AT&T (Brief of SFLC as Amicus Curiae in Support of Petitioner)



No. 05-1056

IN THE
Supreme Court of the United States

MICROSOFT CORPORATION,
Petitioner,
v.
AT&T CORP.,
Respondent.

**On Writ of Certiorari to the
United States Court of Appeals
for the Federal Circuit**

BRIEF OF THE SOFTWARE FREEDOM
LAW CENTER AS *AMICUS CURIAE*
IN SUPPORT OF PETITIONER

EBEN MOGLEN
Counsel of record
DANIEL RAVICHER
RICHARD FONTANA
Software Freedom Law Center
1995 Broadway, 17th Floor
New York, New York 10023
212-461-1900

December 15, 2006

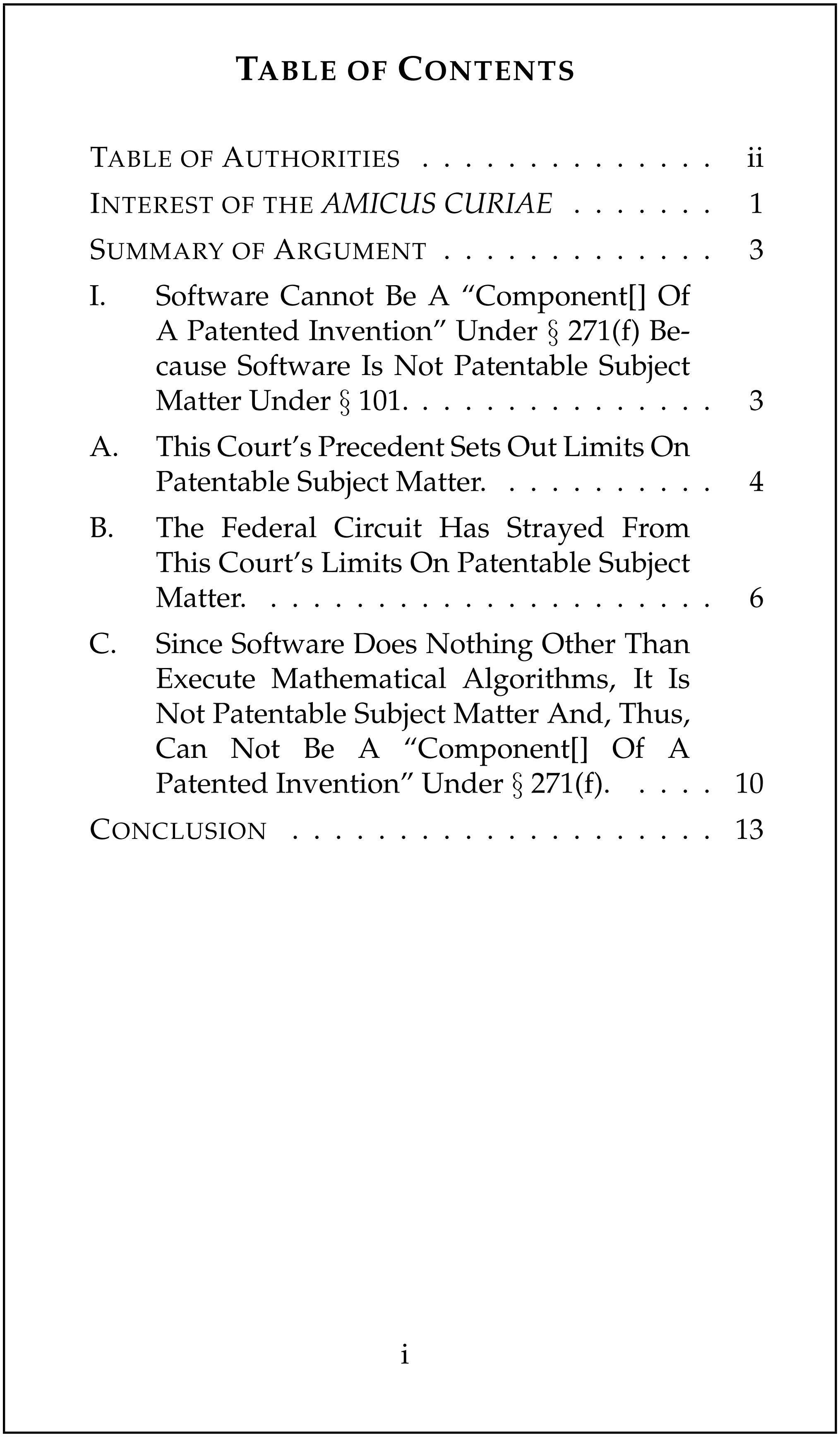


TABLE OF CONTENTS

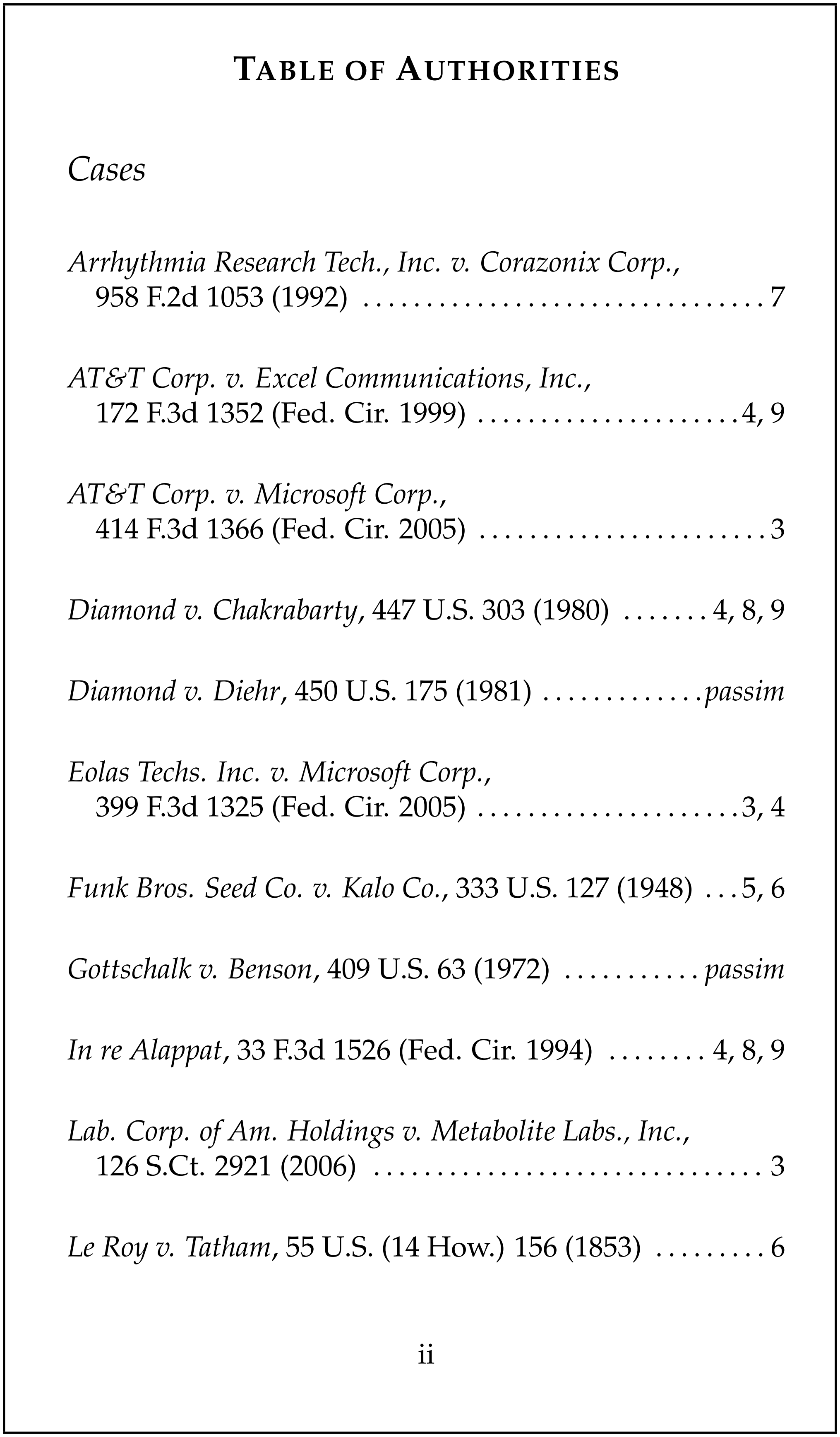


Table of Authorities

Cases

Arrhythmia Research Tech., Inc. v. Corazonix Corp.,
958 F.2d 1053 (1992) 7

AT&T Corp. v. Excel Communications, Inc.,
172 F.3d 1352 (Fed. Cir. 1999) 4, 9

AT&T Corp. v. Microsoft Corp.,
414 F.3d 1366 (Fed. Cir. 2005) 3

Diamond v. Chakrabarty, 447 U.S. 303 (1980) 4, 8, 9

Diamond v. Diehr, 450 U.S. 175 (1981) . . . . *passim*

Eolas Techs. Inc. v. Microsoft Corp.,
399 F.3d 1325 (Fed. Cir. 2005) 3, 4

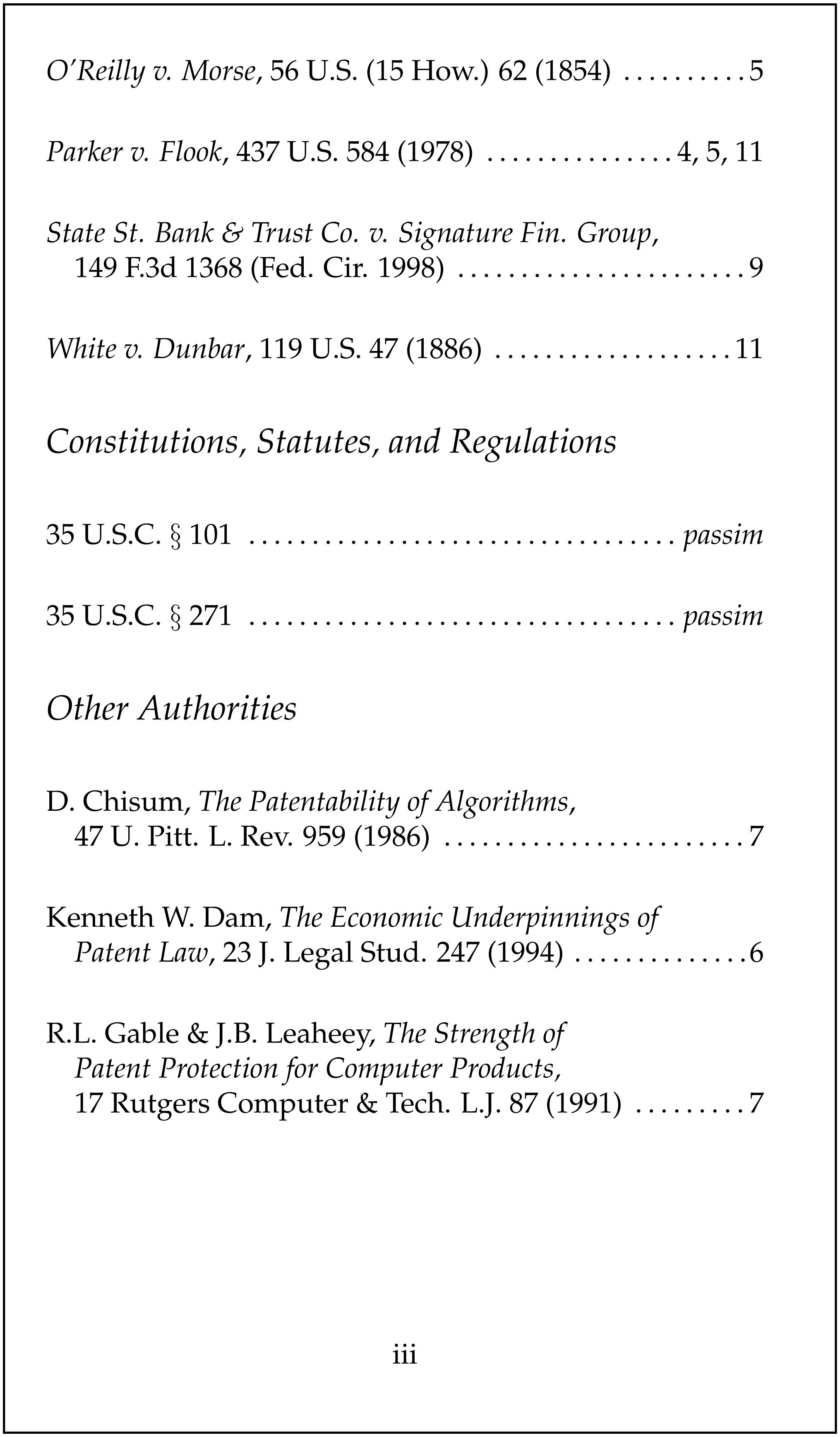
Funk Bros. Seed Co. v. Kalo Co., 333 U.S. 127 (1948) . . . 5, 6

Gottschalk v. Benson, 409 U.S. 63 (1972) . . . . *passim*

In re Alappat, 33 F.3d 1526 (Fed. Cir. 1994) 4, 8, 9

Lab. Corp. of Am. Holdings v. Metabolite Labs., Inc.,
126 S.Ct. 2921 (2006) 3

Le Roy v. Tatham, 55 U.S. (14 How.) 156 (1853) 6



O'Reilly v. Morse, 56 U.S. (15 How.) 62 (1854) 5

Parker v. Flook, 437 U.S. 584 (1978) 4, 5, 11

State St. Bank & Trust Co. v. Signature Fin. Group,
149 F.3d 1368 (Fed. Cir. 1998) 9

White v. Dunbar, 119 U.S. 47 (1886) 11

Constitutions, Statutes, and Regulations

35 U.S.C. § 101 .... *passim*

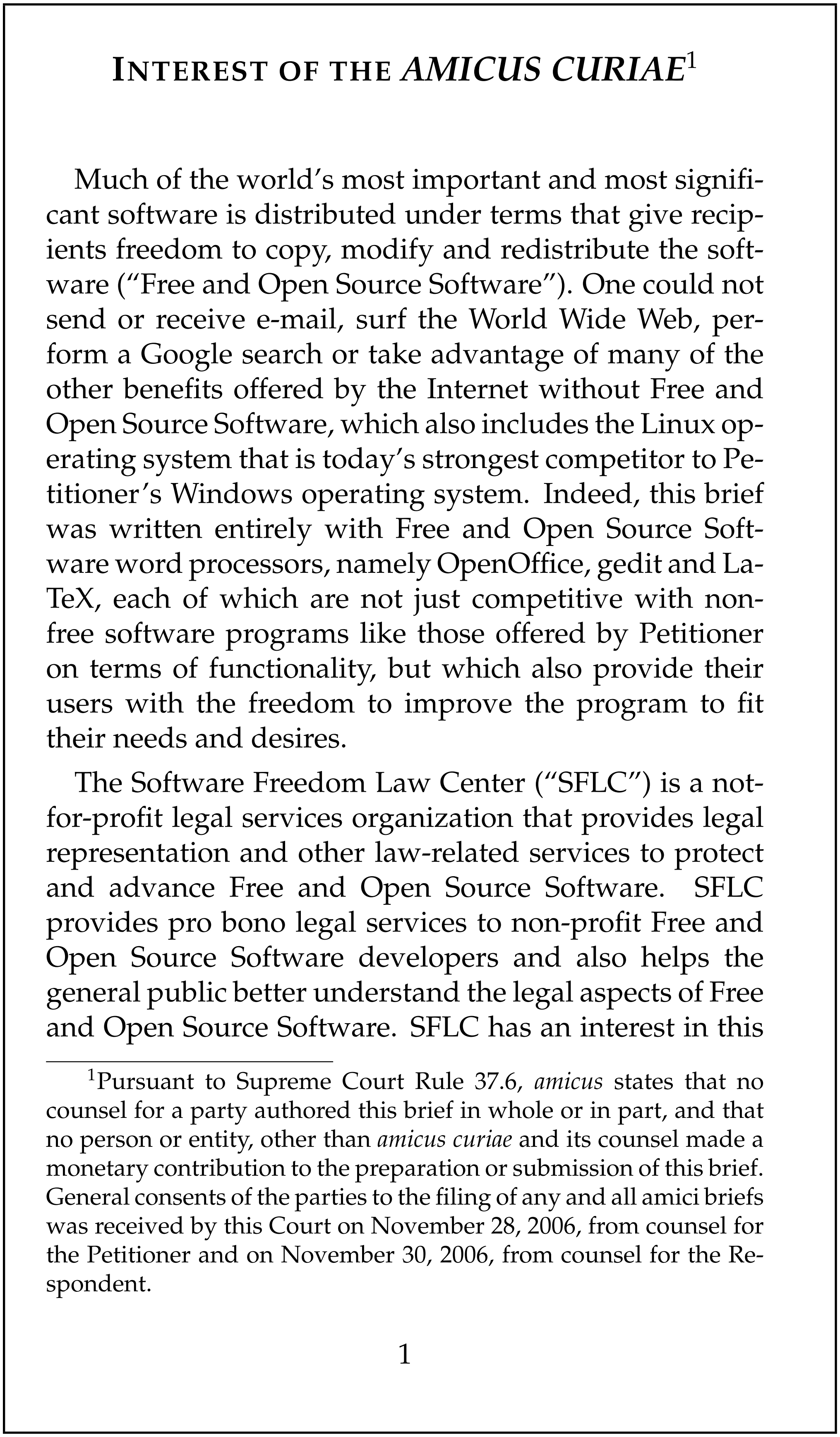
35 U.S.C. § 271 .... *passim*

Other Authorities

D. Chisum, *The Patentability of Algorithms*,
47 U. Pitt. L. Rev. 959 (1986) 7

Kenneth W. Dam, *The Economic Underpinnings of Patent Law*, 23 J. Legal Stud. 247 (1994) 6

R.L. Gable & J.B. Leaheey, *The Strength of Patent Protection for Computer Products*,
17 Rutgers Computer & Tech. L.J. 87 (1991) 7

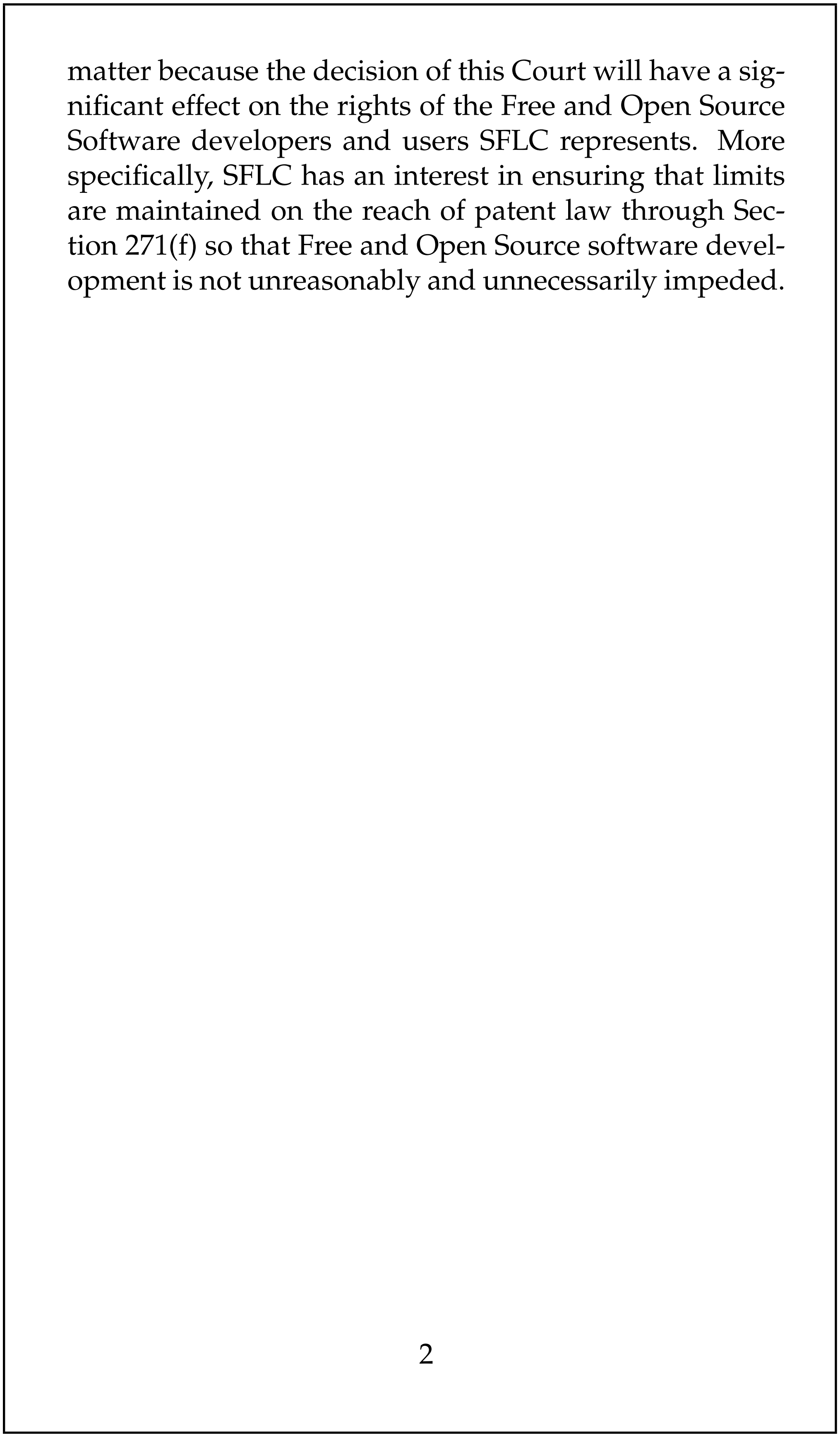


INTEREST OF THE *AMICUS CURIAE*[1]

Much of the world's most important and most significant software is distributed under terms that give recipients freedom to copy, modify and redistribute the software ("Free and Open Source Software"). One could not send or receive e-mail, surf the World Wide Web, perform a Google search or take advantage of many of the other benefits offered by the Internet without Free and Open Source Software, which also includes the Linux operating system that is today's strongest competitor to Petitioner's Windows operating system. Indeed, this brief was written entirely with Free and Open Source Software word processors, namely OpenOffice, gedit and LaTeX, each of which are not just competitive with non-free software programs like those offered by Petitioner on terms of functionality, but which also provide their users with the freedom to improve the program to fit their needs and desires.

The Software Freedom Law Center ("SFLC") is a not-for-profit legal services organization that provides legal representation and other law-related services to protect and advance Free and Open Source Software. SFLC provides pro bono legal services to non-profit Free and Open Source Software developers and also helps the general public better understand the legal aspects of Free and Open Source Software. SFLC has an interest in this

[1] Pursuant to Supreme Court Rule 37.6, *amicus* states that no counsel for a party authored this brief in whole or in part, and that no person or entity, other than *amicus curiae* and its counsel made a monetary contribution to the preparation or submission of this brief. General consents of the parties to the filing of any and all amici briefs was received by this Court on November 28, 2006, from counsel for the Petitioner and on November 30, 2006, from counsel for the Respondent.



matter because the decision of this Court will have a significant effect on the rights of the Free and Open Source Software developers and users SFLC represents. More specifically, SFLC has an interest in ensuring that limits are maintained on the reach of patent law through Section 271(f) so that Free and Open Source software development is not unreasonably and unnecessarily impeded.

SUMMARY OF ARGUMENT

Software can not be a "component[] of a patented invention" under 35 U.S.C. § 271(f) because software is not patentable subject matter under 35 U.S.C. § 101. As such, the Federal Circuit's holding to the contrary in this case is erroneous and should be reversed.

I. Software Cannot Be A "Component[] Of A Patented Invention" Under § 271(f) Because Software Is Not Patentable Subject Matter Under § 101.

The Court of Appeals for the Federal Circuit in this case resolved the issue of whether software may be a "component" of a patented invention under § 271(f) by relying on its contemporaneous *Eolas* decision, which held that "without question, software code alone qualifies as an invention eligible for patenting [under 35 U.S.C. § 101]," and that "every form of invention eligible for patenting [under 35 U.S.C. § 101] falls within the protection of Section 271(f)." *AT&T Corp. v. Microsoft Corp.*, 414 F.3d 1366, 1369 (Fed. Cir. 2005) (*citing Eolas Techs. Inc. v. Microsoft Corp.*, 399 F.3d 1325 (Fed. Cir. 2005)).

While the Federal Circuit is correct that only subject matter eligible for patenting under § 101 can be captured by § 271(f), the Federal Circuit's holding in *Eolas* that software is patentable subject matter conflicts with longstanding precedents of this Court. As noted in the recent opinion of Justice Breyer dissenting from this Court's decision to dismiss as improvidently granted a patentable subject matter challenge, this Court has not approved of the Federal Circuits Section 101 jurisprudence. *See Lab. Corp. of Am. Holdings v. Metabolite Labs., Inc.*, 126 S.Ct.

2921, 2928 (2006) (Breyer, J., *dissenting*) (*discussing* the Federal Circuit's "useful, concrete, and tangible result" test for patentable subject matter and *stating* that "this Court has never made such a statement and, if taken literally, the statement would cover instances where this Court has held the contrary").

To support its holding in *Eolas* that "without question, software code alone qualifies as an invention eligible for patenting," the Federal Circuit relied merely on its own previous decisions. *Eolas*, 399 F.3d at 1339 (*citing In re Alappat*, 33 F.3d 1526 (Fed. Cir. 1994) and *AT&T Corp. v. Excel Communications, Inc.*, 172 F.3d 1352 (Fed. Cir. 1999)). *Eolas* and the earlier cases on which it relied completely ignored this Court's precedent (discussed below) that sets out firm limits on patentable subject matter and that - in fact - excludes software from patentable subject matter.

Therefore, since *Eolas* fails to abide by this Court's precedent regarding patentable subject matter, the Federal Circuit's reliance on *Eolas* for the holding in this case that software can be a "component[] of a patented invention" under § 271(f) is legally erroneous and should be reversed.

A. THIS COURT'S PRECEDENT SETS OUT LIMITS ON PATENTABLE SUBJECT MATTER.

Confronted with the rise of new technologies, this Court has addressed the issue of patentable subject matter several times**.** *Gottschalk v. Benson*, 409 U.S. 63, 71 (1972); *Parker v. Flook*, 437 U.S. 584, 591 (1978); *Diamond v. Chakrabarty*, 447 U.S. 303 (1980); *Diamond v. Diehr*, 450 U.S. 175 (1981). Since before the Civil War, this Court has consistently made it clear that subject matter which would have the practical effect of preempting laws of nature, abstract ideas or mathematical algorithms is ineligible for patent protection. *O'Reilly v. Morse*, 56 U.S. (15 How.) 62, 113 (1854); *Benson*, 409 U.S. at 71. This age-old and time-tested precedent effectively establishes a penumbra of ineligibility for patent protection to safeguard the fundamental policy that laws of nature, abstract ideas and mathematical algorithms be left unrestrained by patents.

This Court stated in *Flook* that to be eligible for patent protection, "[a] process itself, not merely the mathematical algorithm, must be new and useful." 437 U.S. at 591; *Funk Bros. Seed Co. v. Kalo Co.*, 333 U.S. 127, 130 (1948). This Court further stated in *Flook* that it is "incorrect[to] assume[] that if a process application implements a principle in some specific fashion, it automatically falls within the patentable subject matter of § 101." 437 U.S. at 593. This Court explained that such an assumption is based on an impermissibly narrow interpretation of its precedent, including specifically *Benson,* and is "untenable" because "[i]t would make the determination of patentable subject matter depend simply on the draftsman's art and would ill serve the principles underlying the prohibition against patents for 'ideas' or phenomena of nature." *Id.*

In alignment with *Benson* and *Flook*, this Court's decision in *Diehr* held that structures or processes must, when considered as a whole, perform functions intended to be covered by patent law in order to be eligible for patent protection. 450 U.S. at 192. *Diehr* followed and upheld the core holdings of both *Benson* and *Flook*. *Id.* at 190, 191-193 (*citing Benson* and *Flook* repeatedly and *stating* "[o]ur reasoning in *Flook* is in no way inconsistent with our reasoning here").

Benson, *Flook*, *Diehr* and the other decisions of this

Court regarding patentable subject matter consistently established that the inquiry into whether subject matter is eligible for patenting is one of substance, not form. This Court requires that one look, not simply at the language of the patent claim to see if it recites a structure of multiple steps or components, but also at the practical effect of the claim to see if it in fact covers - or otherwise would restrict the public's access to - a principle, law of nature, abstract idea, mathematical formula, mental process, algorithm or other abstract intellectual concept.

This substantive standard ensures that skilled patent draftsmanship is not capable of overcoming one of the core principles of patent law recognized by this Court for more than 150 years that "[a] principle, in the abstract, is a fundamental truth; an original cause; a motive; these cannot be patented, as no one can claim in either of them an exclusive right." *Le Roy v. Tatham*, 55 U.S. (14 How.) 156, 175 (1853); *Funk Bros.*, 333 U.S. at 130; *Benson*, 409 U.S. at 67 ("[p]henomena of nature, though just discovered, mental processes, and abstract intellectual concepts are not patentable, as they are the basic tools of scientific and technological work").

B. The Federal Circuit Has Strayed From This Court's Limits On Patentable Subject Matter.

Many scholars have noted that the creation of the Federal Circuit "did away as a practical matter with Supreme Court jurisdiction in patent cases." Kenneth W. Dam, *The Economic Underpinnings of Patent Law*, 23 J. Legal Stud. 247, 270 (1994). For example, through a series of decisions, the Federal Circuit has abandoned the substantive based standard established by this Court for determining patentable subject matter and replaced it with a more expansive formalistic approach that looks only to see whether a patent claim contains some structure or has some minimal practical utility. The Federal Circuit's form-over-substance approach has come to include virtually anything within patentable subject matter.

Initially, the Federal Circuit used the opinions of legal commentators to justify straying from *Benson* and *Flook*. *Arrhythmia Research Tech., Inc. v. Corazonix Corp.*, 958 F.2d 1053, 1057 n.4 (1992) ("Although commentators have differed in their interpretations of *Benson, Flook,* and *Diehr,* it appears to be *generally agreed* that these decisions represent *evolving views* of the Court, and that the reasoning in *Diehr* not only elaborated on, but in part superseded, that of *Benson* and *Flook*") (emphasis added) (*citing* R.L. Gable & J.B. Leaheey, *The Strength of Patent Protection for Computer Products*, 17 Rutgers Computer & Tech. L.J. 87 (1991); D. Chisum, *The Patentability of Algorithms*, 47 U. Pitt. L. Rev. 959 (1986)). Evidently, the Federal Circuit felt that "general agreement" amongst legal commentators justified abandoning this Court's precedent. In reaching this conclusion, the Federal Circuit also ignored the *Diehr* Court's statement that its decision there was in accord with *Benson* and *Flook*. *Diehr*, 450 U.S. at 185 - 193.

Also in *Arrhythmia*, the Federal Circuit stated that "claims to a specific process or apparatus... will *generally satisfy* section 101." *Id.* at 1058 (emphasis added). This Court's precedent does not, in fact, support the proposition that any process or apparatus "generally satisfies" the requirements of patentable subject matter. *Diehr*, 450 U.S. at 193 ("[a] mathematical formula as such is not accorded the protection of our patent laws... and this principle cannot be circumvented by attempting to limit the use of the formula to a particular technological environment") (citing *Benson* and *Flook*). The new "general rule" promulgated in *Arrhythmia* was a major step in the Federal Circuit’s departure from this Court’s precedent regarding patentable subject matter.

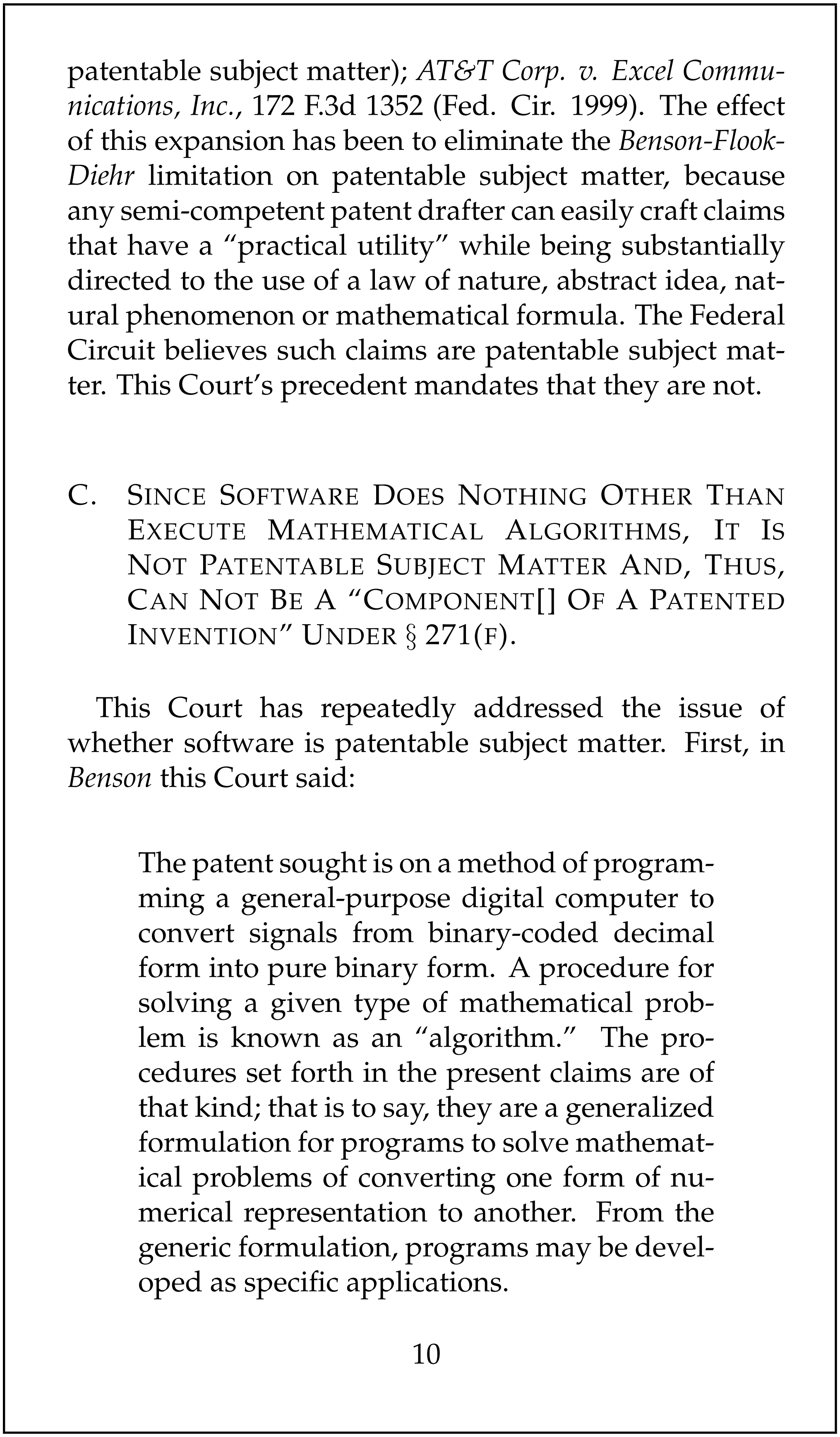
Roughly two years later, the Federal Circuit said that this Court’s precedent on patentable subject matter was too unclear to follow. *In re Alappat*, 33 F.3d 1526, 1543 n.19 and n.20 (Fed. Cir. 1994) (“The Supreme Court has not been clear”, “The Supreme Court has not set forth, however, any consistent or clear explanation“, ”the understandable struggle that the [Supreme] Court was having in articulating a rule”). Contrary to the Federal Circuit’s characterizations, however, this Court’s precedent on patentable subject matter is plainly clear: the analysis is one of substance, not form, and asks whether a patent claim is substantially directed to a law of nature, natural phenomenon, abstract idea or mathematical algorithm.

After disregarding this Court’s precedent as “unclear,” the Federal Circuit substituted its own formalistic approach, which finds that virtually anything is eligible for patenting. *Id*. at 1542 (“[t]he use of the expansive term ’any’ in § 101 represents Congress’s intent not to place any restrictions on the subject matter for which a patent may be obtained”). The Federal Circuit’s approach conflicts with this Court’s precedent. As just one example, it ignores the firm statement in *Diehr* that “[a] mathematical formula does not suddenly become patentable subject matter simply by having the applicant acquiesce to limiting the reach of the patent for the formula to a particular technological use.” 450 U.S. at 193.

In support of its holding, the Federal Circuit cited this Court’s *Chakrabarty* decision for the proposition that, “Congress intended § 101 to extend to ’anything under the sun that is made by man.” *Id.* (citing *Chakrabarty*, 447 U.S. 303, 309). However, the Federal Circuit then went much farther than *Chakrabarty’s* holding by saying, "Thus, it is improper to read into § 101 limitations as to the subject matter that may be patented where the legislative history does not indicate that Congress clearly intended such limitations." *Id.* But such was precisely not this Court's holding in *Chakrabarty*. Immediately following the language quoted by the Federal Circuit, this Court continued to say in *Chakrabarty* that, "[t]his is not to suggest that § 101 has no limits or that it embraces every discovery." 447 U.S. at 309 (emphasis added). In support of that statement, this Court referred to *Flook*, *Benson*, *Funk Bros.* and other cases, and not to any legislative history. Thus, this Court's precedent clearly shows that there are indeed limits on patentable subject matter beyond those expressly stated by Congress. The Federal Circuit's ruling to the contrary was error.

Indeed, *Alappat* was a highly divided *en banc* decision, wherein several members of the Federal Circuit recognized that the majority was making a severe judicial error. *Id.* at 1552, 1562 (Archer, C.J., *dissenting*). Chief Judge Archer said, "Losing sight of the forest for the structure of the trees, the majority today holds that any claim reciting a precise arrangement of structure satisfies 35 U.S.C. §101.... [T]he rationale that leads to this conclusion and the majority's holding that Alappat's rasterizer represents the invention of a machine are illogical, inconsistent with precedent and with sound principles of patent law, and will have untold consequences," and that "the majority's test under § 101 that looks simply to whether specific structure is claimed is [] inconsistent with Supreme Court precedent"). *Id.*

Since *Alappat*, the Federal Circuit has continued its expansion of patentable subject matter through the implementation of its formalistic approach. *State St. Bank & Trust Co. v. Signature Fin. Group*, 149 F.3d 1368 (Fed. Cir. 1998) (holding that anything with a "practical utility" is

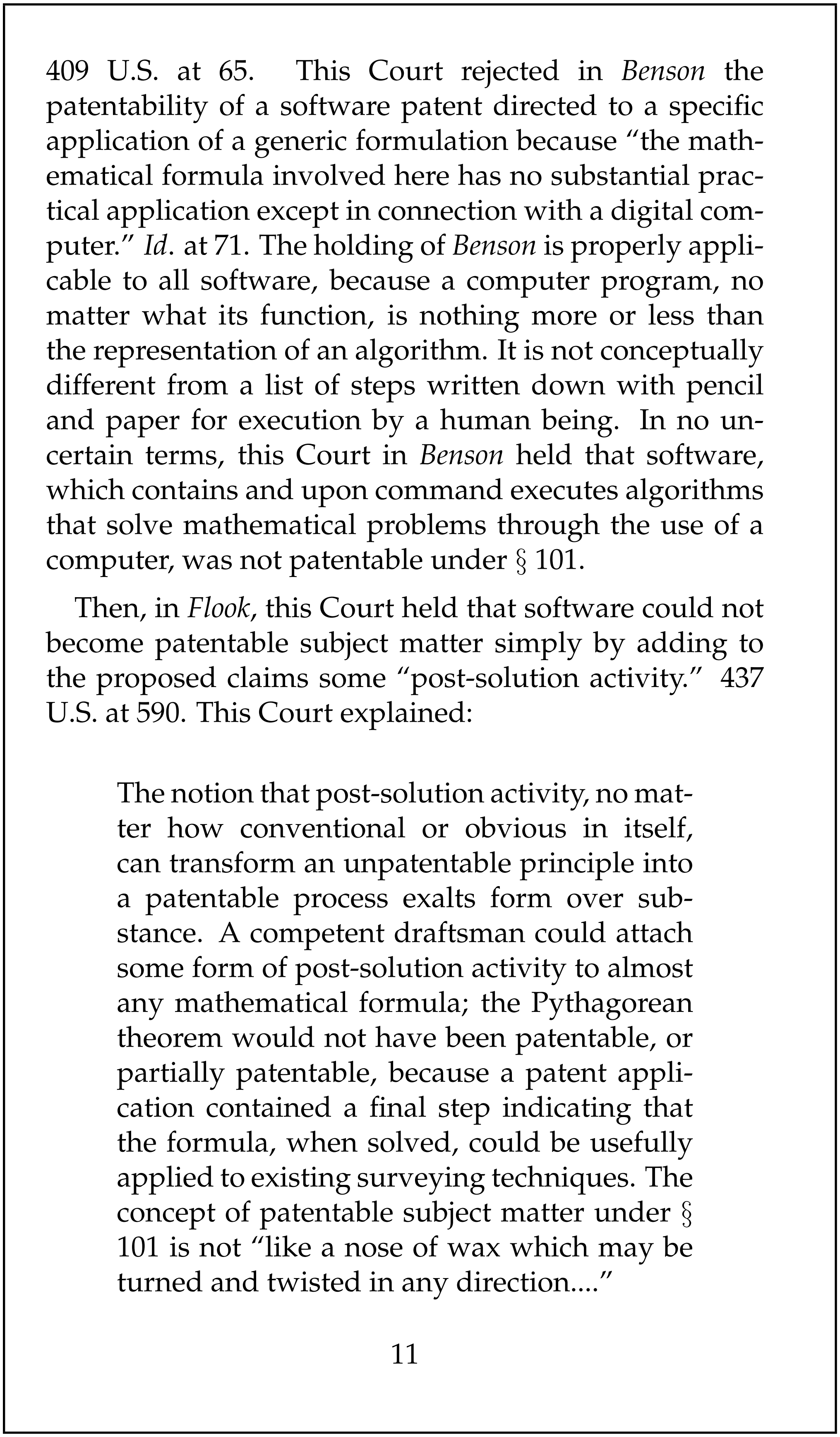


patentable subject matter); *AT&T Corp. v. Excel Communications, Inc.*, 172 F.3d 1352 (Fed. Cir. 1999). The effect of this expansion has been to eliminate the *Benson-Flook-Diehr* limitation on patentable subject matter, because any semi-competent patent drafter can easily craft claims that have a "practical utility" while being substantially directed to the use of a law of nature, abstract idea, natural phenomenon or mathematical formula. The Federal Circuit believes such claims are patentable subject matter. This Court's precedent mandates that they are not.

C. Since Software Does Nothing Other Than Execute Mathematical Algorithms, It Is Not Patentable Subject Matter And, Thus, Can Not Be A "Component[] Of A Patented Invention" Under § 271(f).

This Court has repeatedly addressed the issue of whether software is patentable subject matter. First, in *Benson* this Court said:

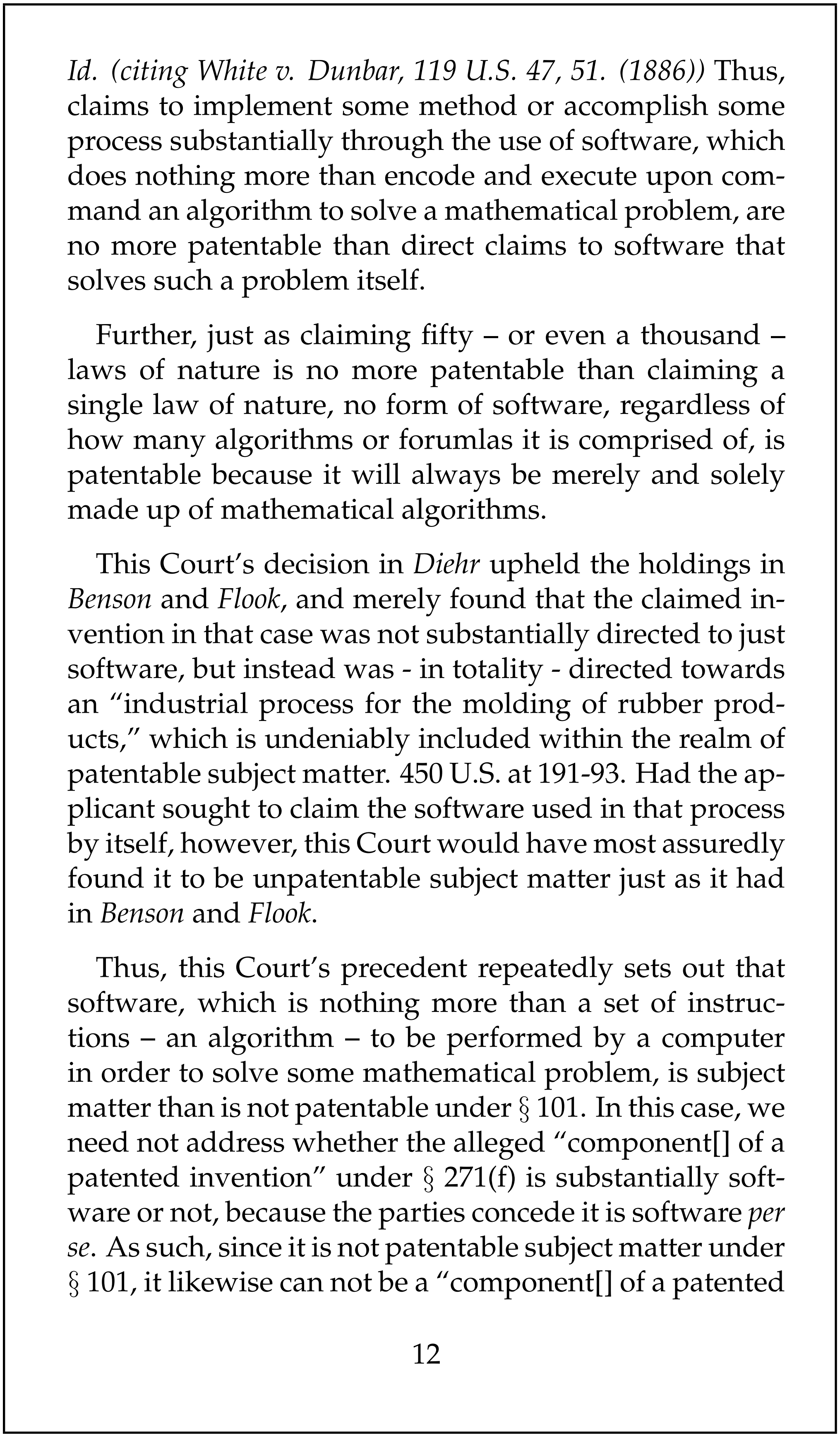
> The patent sought is on a method of programming a general-purpose digital computer to convert signals from binary-coded decimal form into pure binary form. A procedure for solving a given type of mathematical problem is known as an "algorithm." The procedures set forth in the present claims are of that kind; that is to say, they are a generalized formulation for programs to solve mathematical problems of converting one form of numerical representation to another. From the generic formulation, programs may be developed as specific applications.



409 U.S. at 65. This Court rejected in *Benson* the patentability of a software patent directed to a specific application of a generic formulation because "the mathematical formula involved here has no substantial practical application except in connection with a digital computer." *Id*. at 71. The holding of *Benson* is properly applicable to all software, because a computer program, no matter what its function, is nothing more or less than the representation of an algorithm. It is not conceptually different from a list of steps written down with pencil and paper for execution by a human being. In no uncertain terms, this Court in *Benson* held that software, which contains and upon command executes algorithms that solve mathematical problems through the use of a computer, was not patentable under § 101.

Then, in *Flook*, this Court held that software could not become patentable subject matter simply by adding to the proposed claims some "post-solution activity." 437 U.S. at 590. This Court explained:

> The notion that post-solution activity, no matter how conventional or obvious in itself, can transform an unpatentable principle into a patentable process exalts form over substance. A competent draftsman could attach some form of post-solution activity to almost any mathematical formula; the Pythagorean theorem would not have been patentable, or partially patentable, because a patent application contained a final step indicating that the formula, when solved, could be usefully applied to existing surveying techniques. The concept of patentable subject matter under § 101 is not "like a nose of wax which may be turned and twisted in any direction...."



Id. (citing White v. Dunbar, 119 U.S. 47, 51. (1886)) Thus, claims to implement some method or accomplish some process substantially through the use of software, which does nothing more than encode and execute upon command an algorithm to solve a mathematical problem, are no more patentable than direct claims to software that solves such a problem itself.

Further, just as claiming fifty – or even a thousand – laws of nature is no more patentable than claiming a single law of nature, no form of software, regardless of how many algorithms or forumlas it is comprised of, is patentable because it will always be merely and solely made up of mathematical algorithms.

This Court's decision in *Diehr* upheld the holdings in *Benson* and *Flook*, and merely found that the claimed invention in that case was not substantially directed to just software, but instead was - in totality - directed towards an "industrial process for the molding of rubber products," which is undeniably included within the realm of patentable subject matter. 450 U.S. at 191-93. Had the applicant sought to claim the software used in that process by itself, however, this Court would have most assuredly found it to be unpatentable subject matter just as it had in *Benson* and *Flook*.

Thus, this Court's precedent repeatedly sets out that software, which is nothing more than a set of instructions – an algorithm – to be performed by a computer in order to solve some mathematical problem, is subject matter than is not patentable under § 101. In this case, we need not address whether the alleged "component[] of a patented invention" under § 271(f) is substantially software or not, because the parties concede it is software *per se*. As such, since it is not patentable subject matter under § 101, it likewise can not be a "component[] of a patented invention" under § 271(f) and the Federal Circuit's holding in this case to the contrary was judicial error.

CONCLUSION

For the foregoing reasons, this Court should reverse the Federal Circuit's decision.

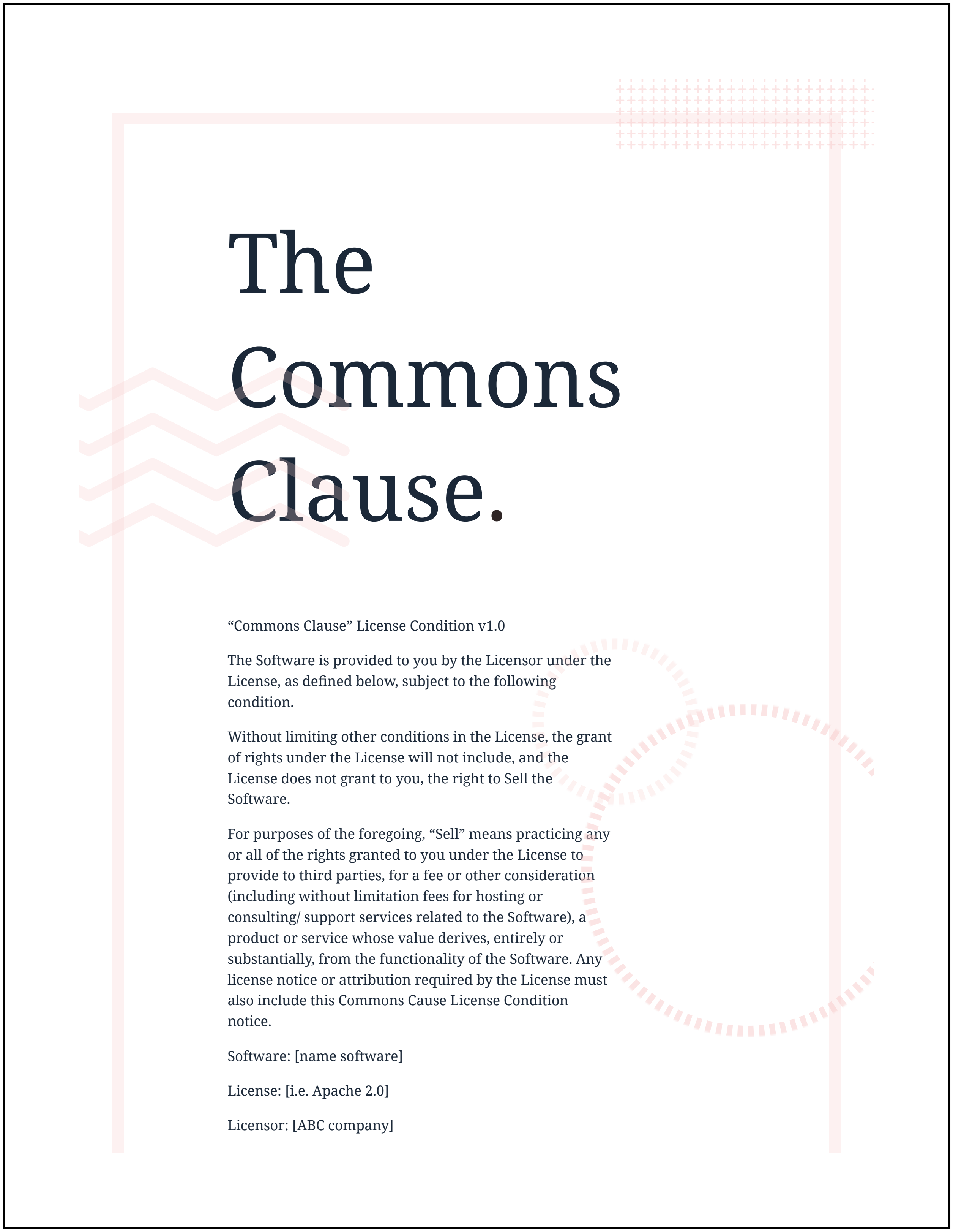
Respectfully submitted.

EBEN MOGLEN
Counsel of record
DANIEL RAVICHER
RICHARD FONTANA
Software Freedom Law Center
1995 Broadway, 17th Floor
New York, New York 10023
212-461-1900

December 15, 2006

Chapter 9

The Commons Clause and FAQ



The Commons Clause.

“Commons Clause” License Condition v1.0

The Software is provided to you by the Licensor under the License, as defined below, subject to the following condition.

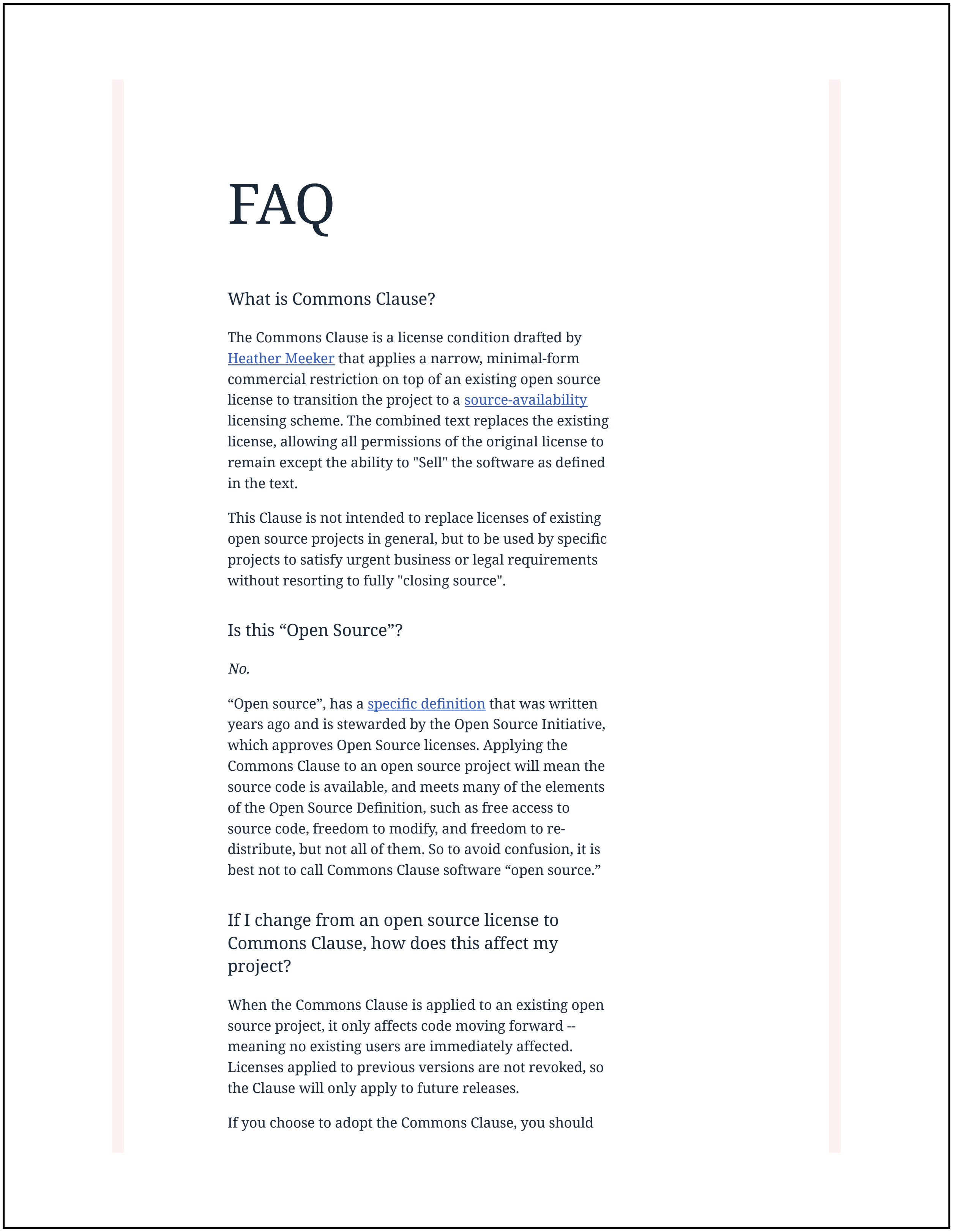
Without limiting other conditions in the License, the grant of rights under the License will not include, and the License does not grant to you, the right to Sell the Software.

For purposes of the foregoing, “Sell” means practicing any or all of the rights granted to you under the License to provide to third parties, for a fee or other consideration (including without limitation fees for hosting or consulting/ support services related to the Software), a product or service whose value derives, entirely or substantially, from the functionality of the Software. Any license notice or attribution required by the License must also include this Commons Cause License Condition notice.

Software: [name software]

License: [i.e. Apache 2.0]

Licensor: [ABC company]



FAQ

What is Commons Clause?

The Commons Clause is a license condition drafted by Heather Meeker that applies a narrow, minimal-form commercial restriction on top of an existing open source license to transition the project to a source-availability licensing scheme. The combined text replaces the existing license, allowing all permissions of the original license to remain except the ability to "Sell" the software as defined in the text.

This Clause is not intended to replace licenses of existing open source projects in general, but to be used by specific projects to satisfy urgent business or legal requirements without resorting to fully "closing source".

Is this “Open Source”?

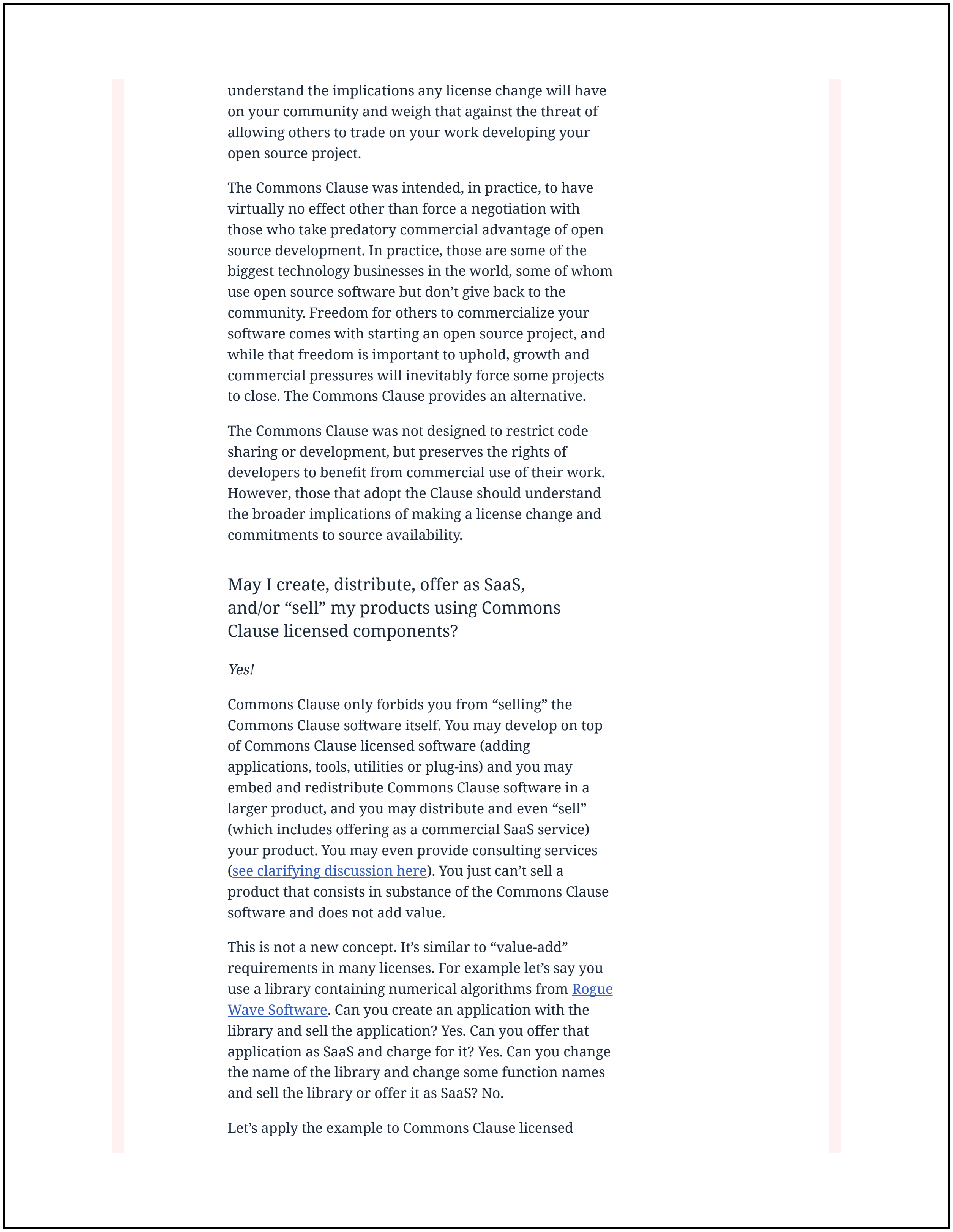
No.

“Open source”, has a specific definition that was written years ago and is stewarded by the Open Source Initiative, which approves Open Source licenses. Applying the Commons Clause to an open source project will mean the source code is available, and meets many of the elements of the Open Source Definition, such as free access to source code, freedom to modify, and freedom to re-distribute, but not all of them. So to avoid confusion, it is best not to call Commons Clause software “open source.”

If I change from an open source license to Commons Clause, how does this affect my project?

When the Commons Clause is applied to an existing open source project, it only affects code moving forward -- meaning no existing users are immediately affected. Licenses applied to previous versions are not revoked, so the Clause will only apply to future releases.

If you choose to adopt the Commons Clause, you should



understand the implications any license change will have on your community and weigh that against the threat of allowing others to trade on your work developing your open source project.

The Commons Clause was intended, in practice, to have virtually no effect other than force a negotiation with those who take predatory commercial advantage of open source development. In practice, those are some of the biggest technology businesses in the world, some of whom use open source software but don’t give back to the community. Freedom for others to commercialize your software comes with starting an open source project, and while that freedom is important to uphold, growth and commercial pressures will inevitably force some projects to close. The Commons Clause provides an alternative.

The Commons Clause was not designed to restrict code sharing or development, but preserves the rights of developers to benefit from commercial use of their work. However, those that adopt the Clause should understand the broader implications of making a license change and commitments to source availability.

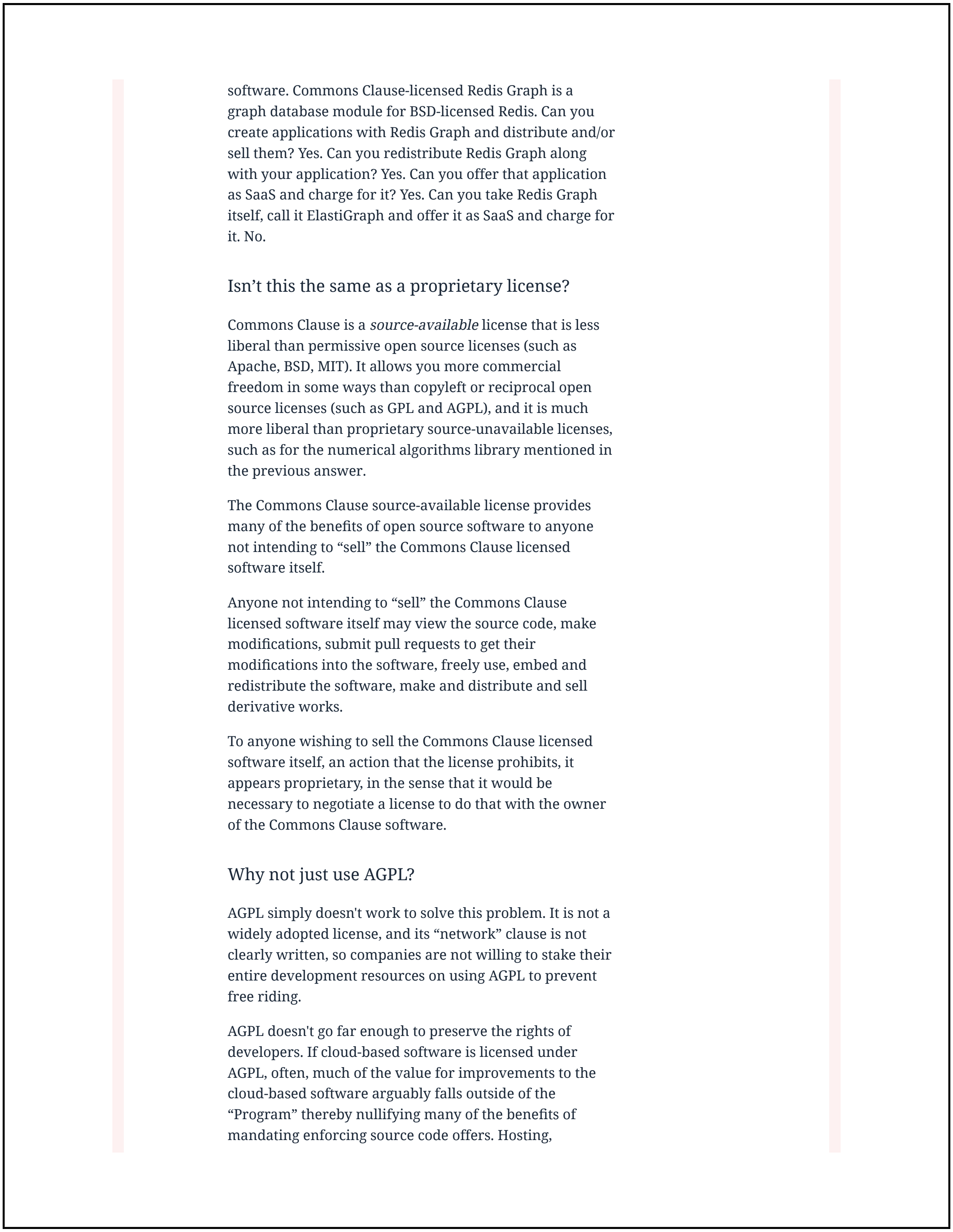
May I create, distribute, offer as SaaS, and/or “sell” my products using Commons Clause licensed components?

Yes!

Commons Clause only forbids you from “selling” the Commons Clause software itself. You may develop on top of Commons Clause licensed software (adding applications, tools, utilities or plug-ins) and you may embed and redistribute Commons Clause software in a larger product, and you may distribute and even “sell” (which includes offering as a commercial SaaS service) your product. You may even provide consulting services (see clarifying discussion here). You just can’t sell a product that consists in substance of the Commons Clause software and does not add value.

This is not a new concept. It’s similar to “value-add” requirements in many licenses. For example let’s say you use a library containing numerical algorithms from Rogue Wave Software. Can you create an application with the library and sell the application? Yes. Can you offer that application as SaaS and charge for it? Yes. Can you change the name of the library and change some function names and sell the library or offer it as SaaS? No.

Let’s apply the example to Commons Clause licensed



software. Commons Clause-licensed Redis Graph is a graph database module for BSD-licensed Redis. Can you create applications with Redis Graph and distribute and/or sell them? Yes. Can you redistribute Redis Graph along with your application? Yes. Can you offer that application as SaaS and charge for it? Yes. Can you take Redis Graph itself, call it ElastiGraph and offer it as SaaS and charge for it. No.

Isn’t this the same as a proprietary license?

Commons Clause is a *source-available* license that is less liberal than permissive open source licenses (such as Apache, BSD, MIT). It allows you more commercial freedom in some ways than copyleft or reciprocal open source licenses (such as GPL and AGPL), and it is much more liberal than proprietary source-unavailable licenses, such as for the numerical algorithms library mentioned in the previous answer.

The Commons Clause source-available license provides many of the benefits of open source software to anyone not intending to “sell” the Commons Clause licensed software itself.

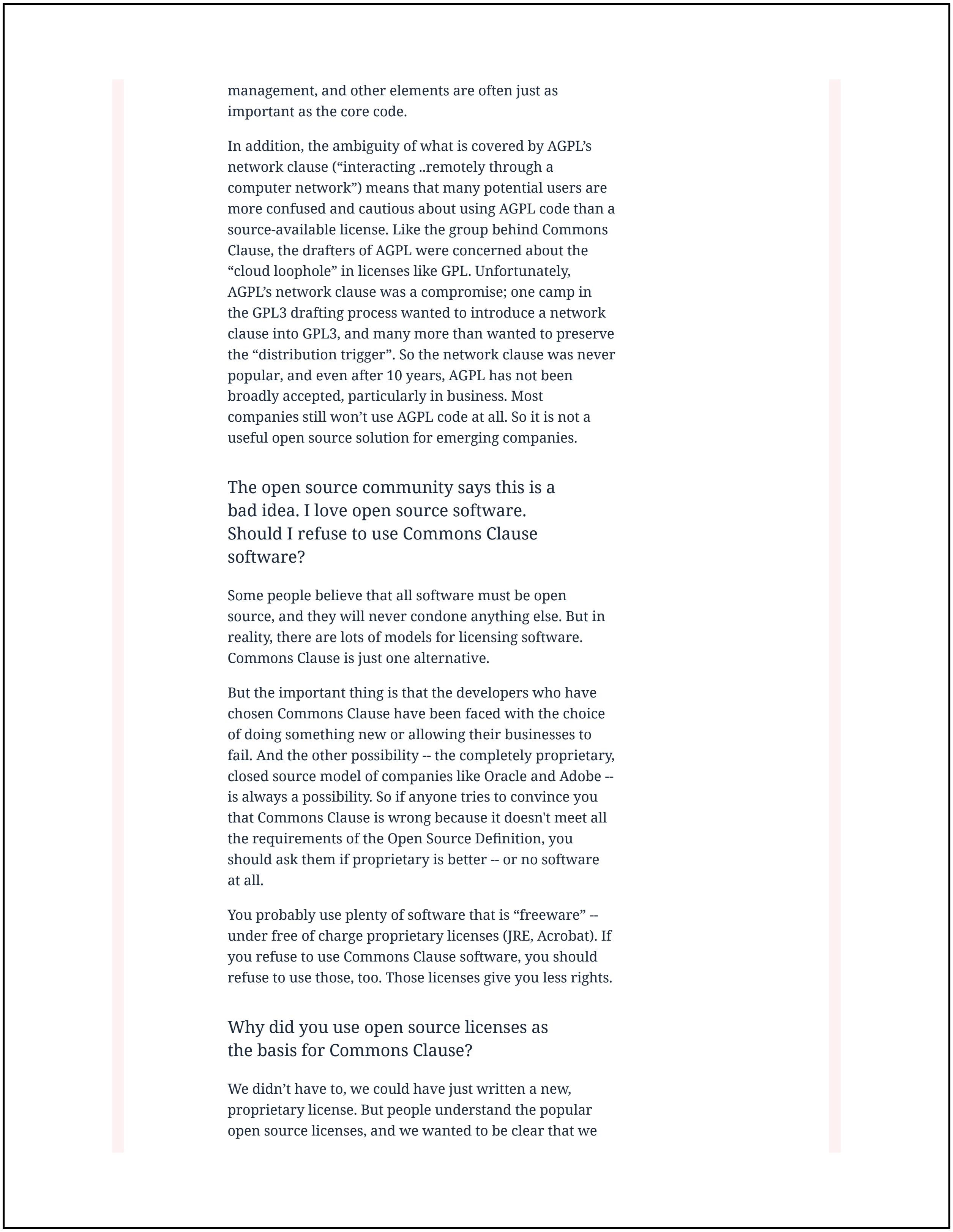
Anyone not intending to “sell” the Commons Clause licensed software itself may view the source code, make modifications, submit pull requests to get their modifications into the software, freely use, embed and redistribute the software, make and distribute and sell derivative works.

To anyone wishing to sell the Commons Clause licensed software itself, an action that the license prohibits, it appears proprietary, in the sense that it would be necessary to negotiate a license to do that with the owner of the Commons Clause software.

Why not just use AGPL?

AGPL simply doesn't work to solve this problem. It is not a widely adopted license, and its “network” clause is not clearly written, so companies are not willing to stake their entire development resources on using AGPL to prevent free riding.

AGPL doesn't go far enough to preserve the rights of developers. If cloud-based software is licensed under AGPL, often, much of the value for improvements to the cloud-based software arguably falls outside of the “Program” thereby nullifying many of the benefits of mandating enforcing source code offers. Hosting,



management, and other elements are often just as important as the core code.

In addition, the ambiguity of what is covered by AGPL's network clause ("interacting ..remotely through a computer network") means that many potential users are more confused and cautious about using AGPL code than a source-available license. Like the group behind Commons Clause, the drafters of AGPL were concerned about the "cloud loophole" in licenses like GPL. Unfortunately, AGPL's network clause was a compromise; one camp in the GPL3 drafting process wanted to introduce a network clause into GPL3, and many more than wanted to preserve the "distribution trigger". So the network clause was never popular, and even after 10 years, AGPL has not been broadly accepted, particularly in business. Most companies still won't use AGPL code at all. So it is not a useful open source solution for emerging companies.

The open source community says this is a bad idea. I love open source software. Should I refuse to use Commons Clause software?

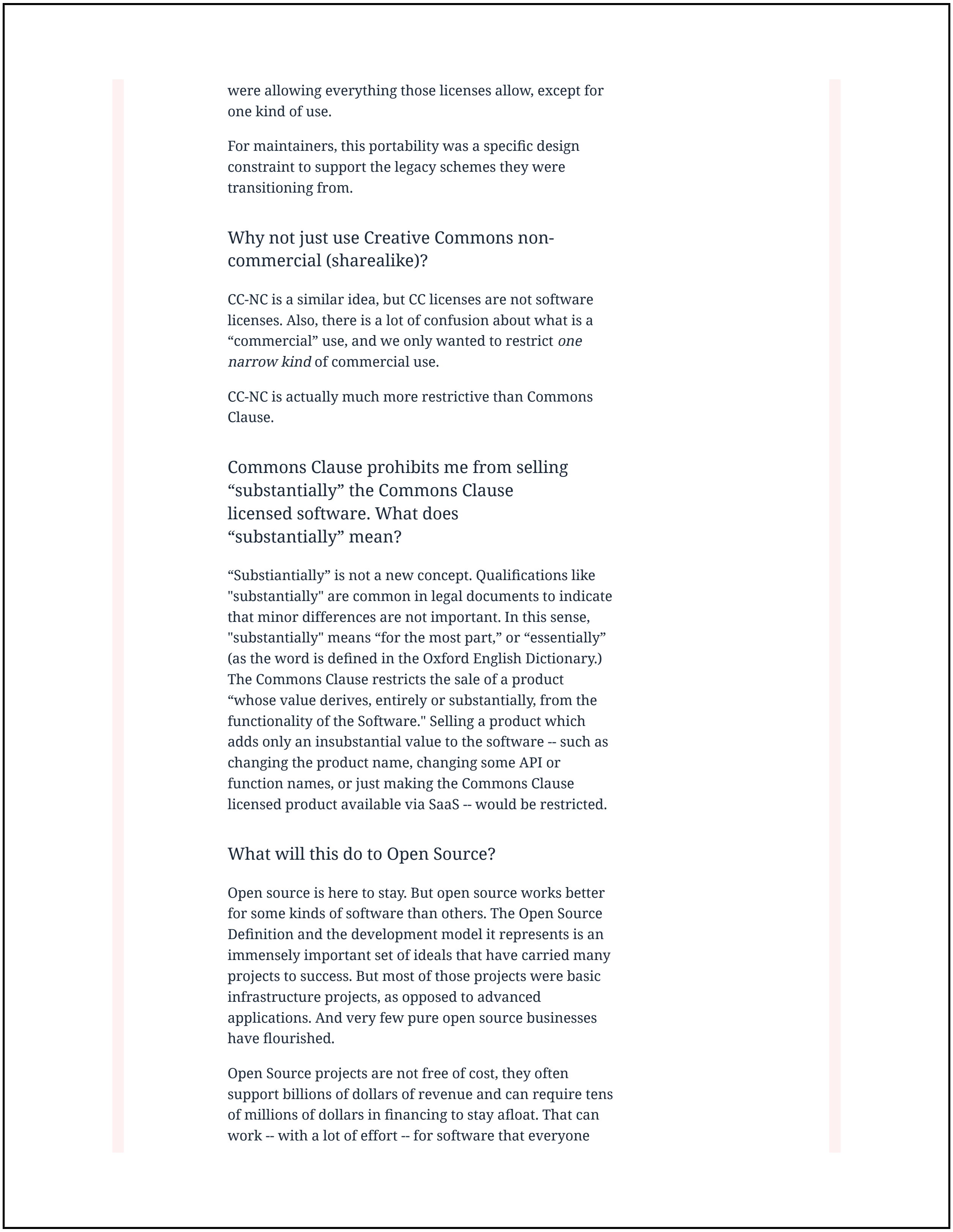
Some people believe that all software must be open source, and they will never condone anything else. But in reality, there are lots of models for licensing software. Commons Clause is just one alternative.

But the important thing is that the developers who have chosen Commons Clause have been faced with the choice of doing something new or allowing their businesses to fail. And the other possibility -- the completely proprietary, closed source model of companies like Oracle and Adobe -- is always a possibility. So if anyone tries to convince you that Commons Clause is wrong because it doesn't meet all the requirements of the Open Source Definition, you should ask them if proprietary is better -- or no software at all.

You probably use plenty of software that is "freeware" -- under free of charge proprietary licenses (JRE, Acrobat). If you refuse to use Commons Clause software, you should refuse to use those, too. Those licenses give you less rights.

Why did you use open source licenses as the basis for Commons Clause?

We didn't have to, we could have just written a new, proprietary license. But people understand the popular open source licenses, and we wanted to be clear that we



were allowing everything those licenses allow, except for one kind of use.

For maintainers, this portability was a specific design constraint to support the legacy schemes they were transitioning from.

Why not just use Creative Commons non-commercial (sharealike)?

CC-NC is a similar idea, but CC licenses are not software licenses. Also, there is a lot of confusion about what is a “commercial” use, and we only wanted to restrict *one narrow kind* of commercial use.

CC-NC is actually much more restrictive than Commons Clause.

Commons Clause prohibits me from selling “substantially” the Commons Clause licensed software. What does “substantially” mean?

“Substiantially” is not a new concept. Qualifications like "substantially" are common in legal documents to indicate that minor differences are not important. In this sense, "substantially" means “for the most part,” or “essentially” (as the word is defined in the Oxford English Dictionary.) The Commons Clause restricts the sale of a product “whose value derives, entirely or substantially, from the functionality of the Software." Selling a product which adds only an insubstantial value to the software -- such as changing the product name, changing some API or function names, or just making the Commons Clause licensed product available via SaaS -- would be restricted.

What will this do to Open Source?

Open source is here to stay. But open source works better for some kinds of software than others. The Open Source Definition and the development model it represents is an immensely important set of ideals that have carried many projects to success. But most of those projects were basic infrastructure projects, as opposed to advanced applications. And very few pure open source businesses have flourished.

Open Source projects are not free of cost, they often support billions of dollars of revenue and can require tens of millions of dollars in financing to stay afloat. That can work -- with a lot of effort -- for software that everyone

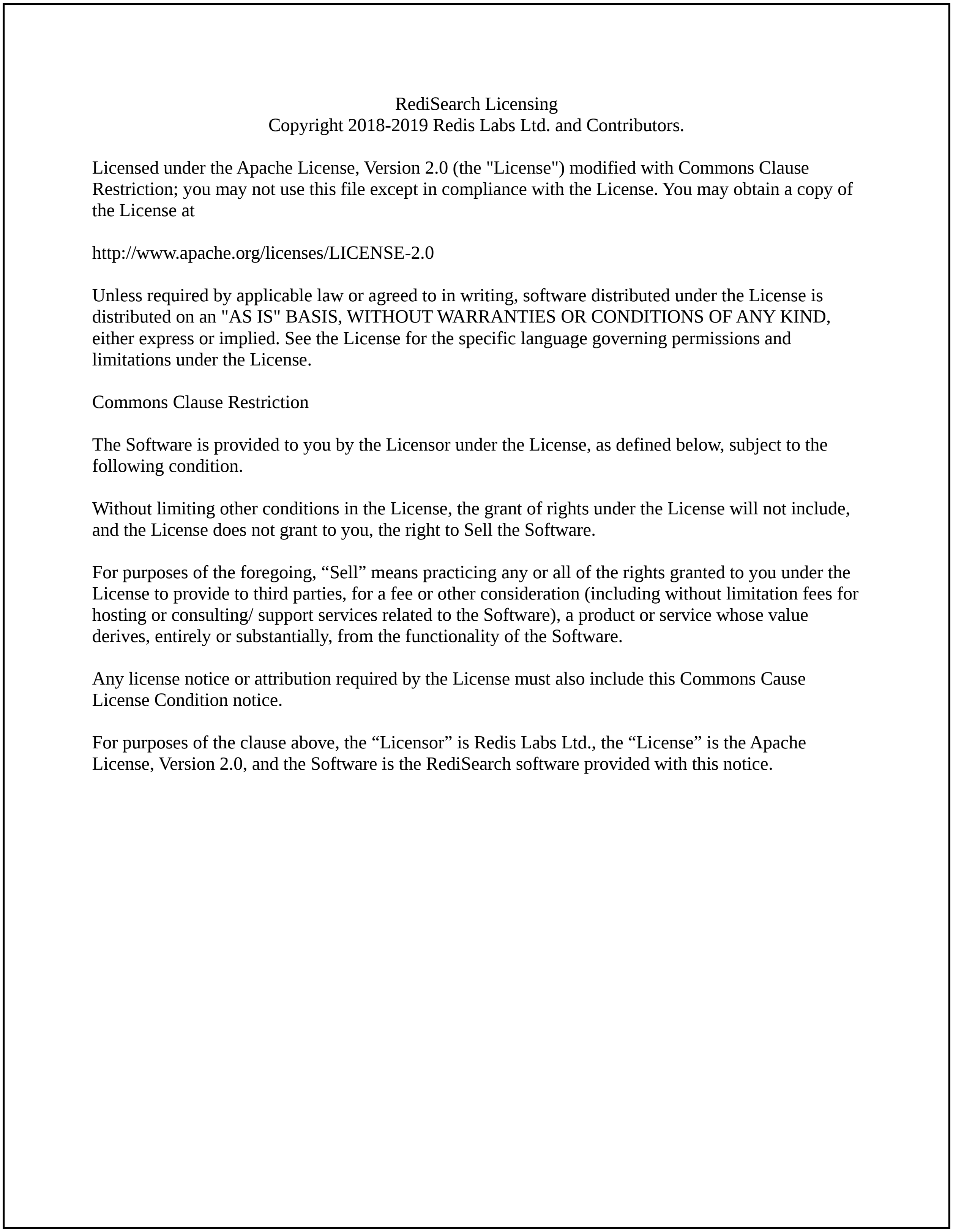
uses, like operating systems. Also, lots of companies are successful using open source -- when they are selling something else, like hardware or services or dual-license upsell modules. But many software companies can’t keep the doors open with an open source licensing model.

The Commons Clause was drafted by a group of developers behind many of the world’s most popular open source projects who feel a lot of pain and pressure from a rapidly-developing business ecosystem and the realities of the cost of developing projects. It wasn’t created to end open source, but start a conversation on what we can do to meet the financial needs of commercial software projects and the communities behind them.

Commons Clause · Contributed by FOSSA (@getfossa) · View Source on GitHub · Website Design by Roka

Chapter 10

RediSearch Licensing



RediSearch Licensing
Copyright 2018-2019 Redis Labs Ltd. and Contributors.

Licensed under the Apache License, Version 2.0 (the "License") modified with Commons Clause Restriction; you may not use this file except in compliance with the License. You may obtain a copy of the License at

http://www.apache.org/licenses/LICENSE-2.0

Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

Commons Clause Restriction

The Software is provided to you by the Licensor under the License, as defined below, subject to the following condition.

Without limiting other conditions in the License, the grant of rights under the License will not include, and the License does not grant to you, the right to Sell the Software.

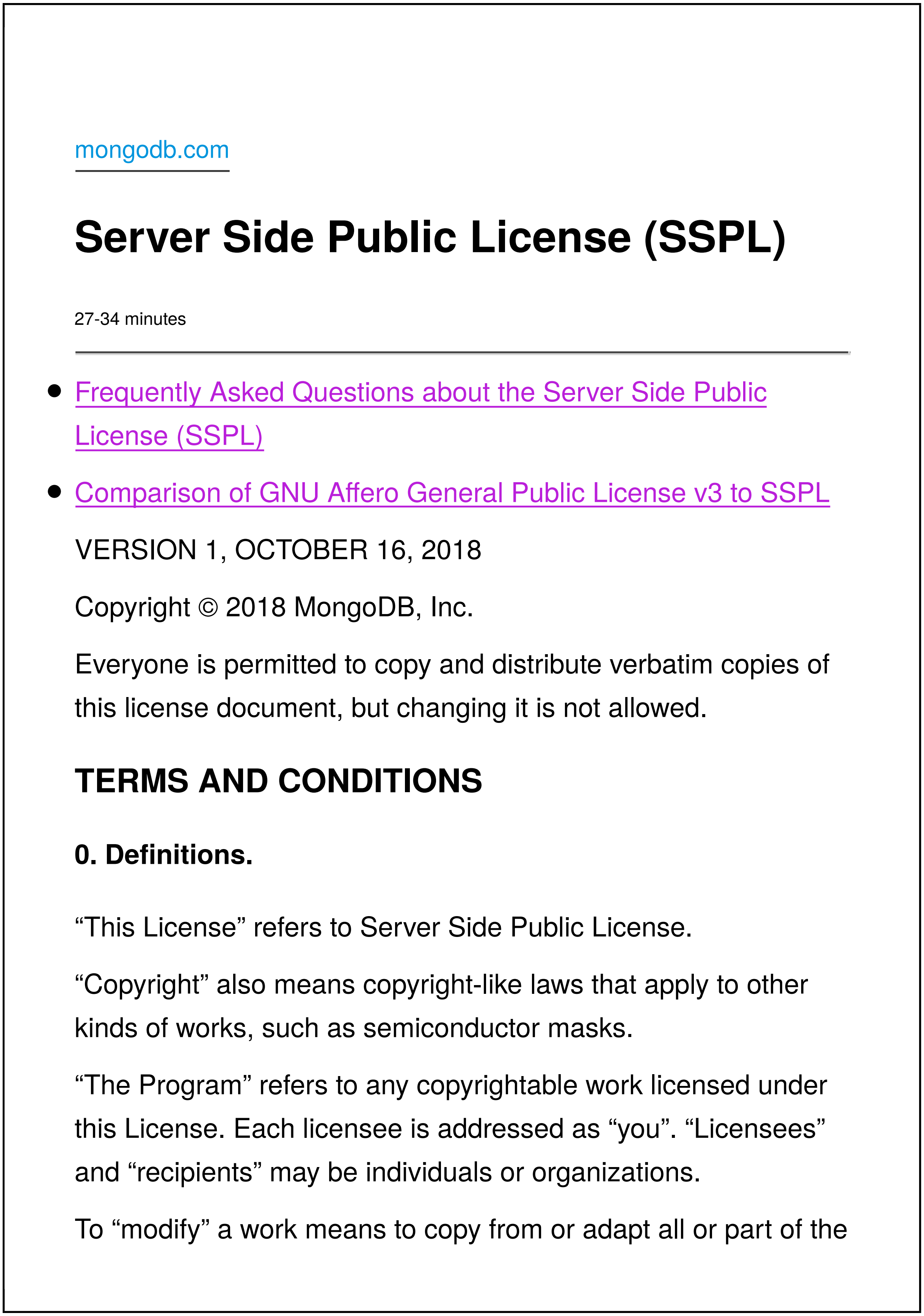
For purposes of the foregoing, “Sell” means practicing any or all of the rights granted to you under the License to provide to third parties, for a fee or other consideration (including without limitation fees for hosting or consulting/ support services related to the Software), a product or service whose value derives, entirely or substantially, from the functionality of the Software.

Any license notice or attribution required by the License must also include this Commons Cause License Condition notice.

For purposes of the clause above, the “Licensor” is Redis Labs Ltd., the “License” is the Apache License, Version 2.0, and the Software is the RediSearch software provided with this notice.

Chapter 11

Mongo DB Server Side Public License (SSPL)



mongodb.com

Server Side Public License (SSPL)

27-34 minutes

- Frequently Asked Questions about the Server Side Public License (SSPL)
- Comparison of GNU Affero General Public License v3 to SSPL

VERSION 1, OCTOBER 16, 2018

Copyright © 2018 MongoDB, Inc.

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

TERMS AND CONDITIONS

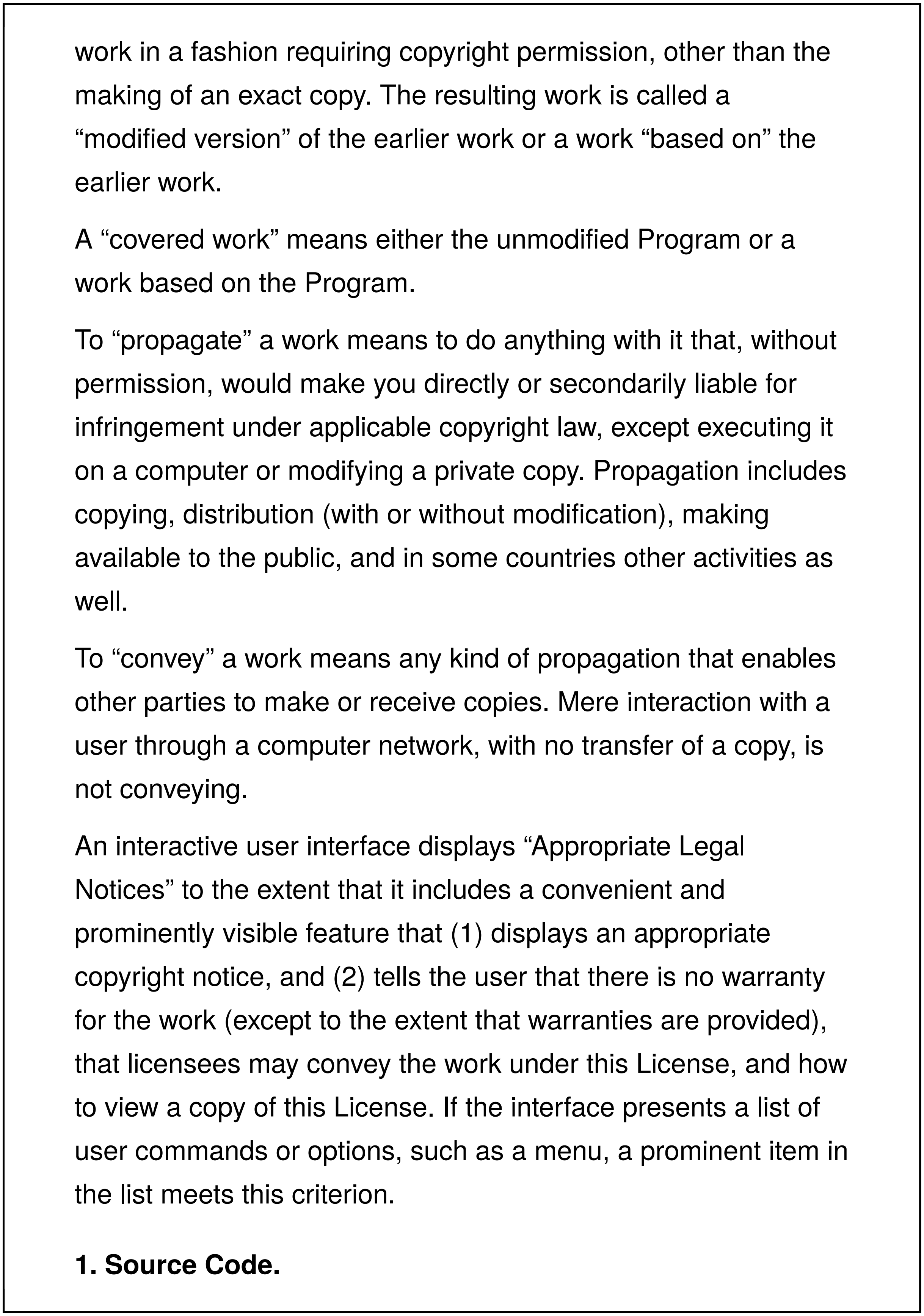
0. Definitions.

“This License” refers to Server Side Public License.

“Copyright” also means copyright-like laws that apply to other kinds of works, such as semiconductor masks.

“The Program” refers to any copyrightable work licensed under this License. Each licensee is addressed as “you”. “Licensees” and “recipients” may be individuals or organizations.

To “modify” a work means to copy from or adapt all or part of the



work in a fashion requiring copyright permission, other than the making of an exact copy. The resulting work is called a “modified version” of the earlier work or a work “based on” the earlier work.

A “covered work” means either the unmodified Program or a work based on the Program.

To “propagate” a work means to do anything with it that, without permission, would make you directly or secondarily liable for infringement under applicable copyright law, except executing it on a computer or modifying a private copy. Propagation includes copying, distribution (with or without modification), making available to the public, and in some countries other activities as well.

To “convey” a work means any kind of propagation that enables other parties to make or receive copies. Mere interaction with a user through a computer network, with no transfer of a copy, is not conveying.

An interactive user interface displays “Appropriate Legal Notices” to the extent that it includes a convenient and prominently visible feature that (1) displays an appropriate copyright notice, and (2) tells the user that there is no warranty for the work (except to the extent that warranties are provided), that licensees may convey the work under this License, and how to view a copy of this License. If the interface presents a list of user commands or options, such as a menu, a prominent item in the list meets this criterion.

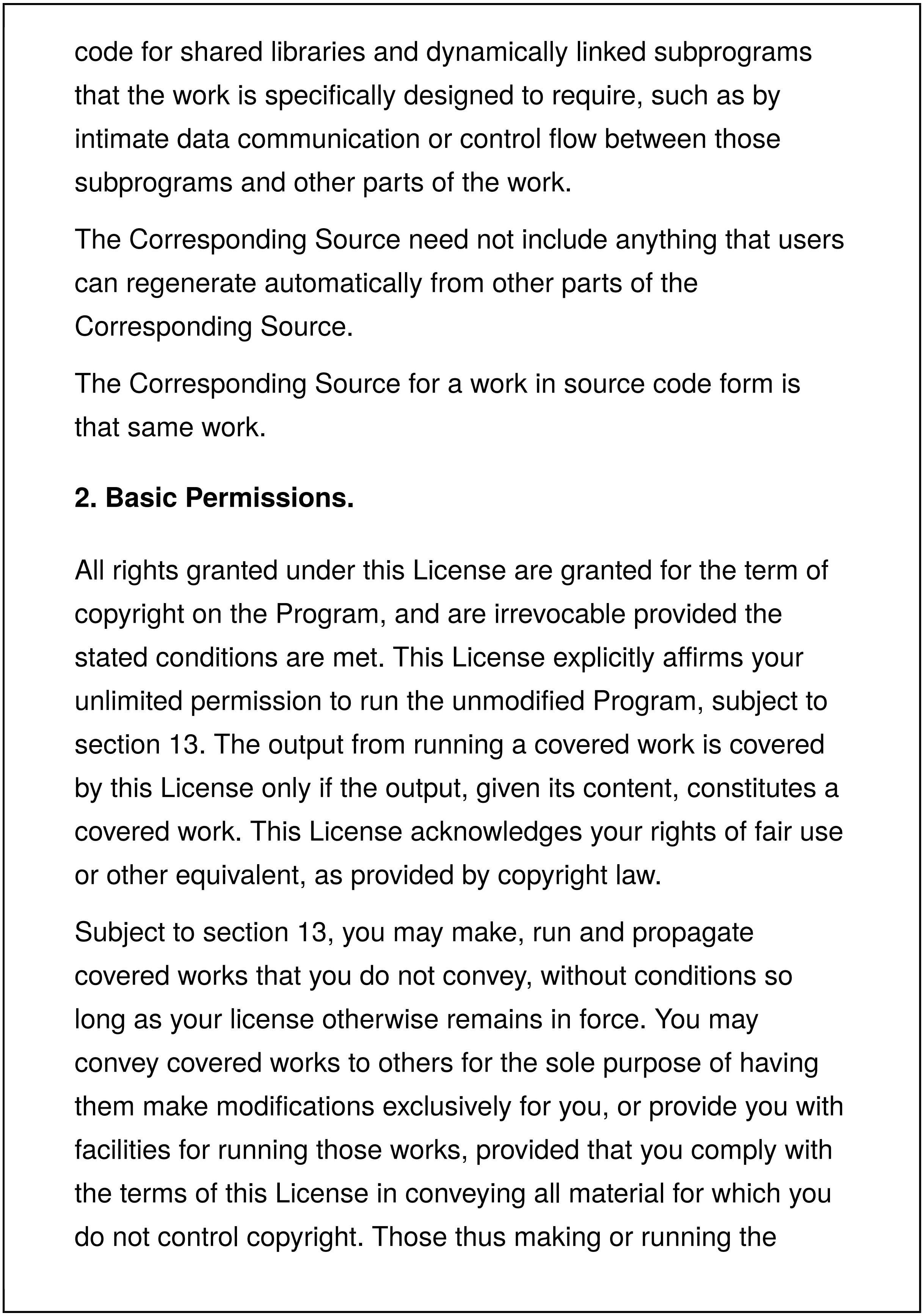
1. Source Code.

The “source code” for a work means the preferred form of the work for making modifications to it. “Object code” means any non-source form of a work.

A “Standard Interface” means an interface that either is an official standard defined by a recognized standards body, or, in the case of interfaces specified for a particular programming language, one that is widely used among developers working in that language.

The “System Libraries” of an executable work include anything, other than the work as a whole, that (a) is included in the normal form of packaging a Major Component, but which is not part of that Major Component, and (b) serves only to enable use of the work with that Major Component, or to implement a Standard Interface for which an implementation is available to the public in source code form. A “Major Component”, in this context, means a major essential component (kernel, window system, and so on) of the specific operating system (if any) on which the executable work runs, or a compiler used to produce the work, or an object code interpreter used to run it.

The “Corresponding Source” for a work in object code form means all the source code needed to generate, install, and (for an executable work) run the object code and to modify the work, including scripts to control those activities. However, it does not include the work's System Libraries, or general-purpose tools or generally available free programs which are used unmodified in performing those activities but which are not part of the work. For example, Corresponding Source includes interface definition files associated with source files for the work, and the source



code for shared libraries and dynamically linked subprograms that the work is specifically designed to require, such as by intimate data communication or control flow between those subprograms and other parts of the work.

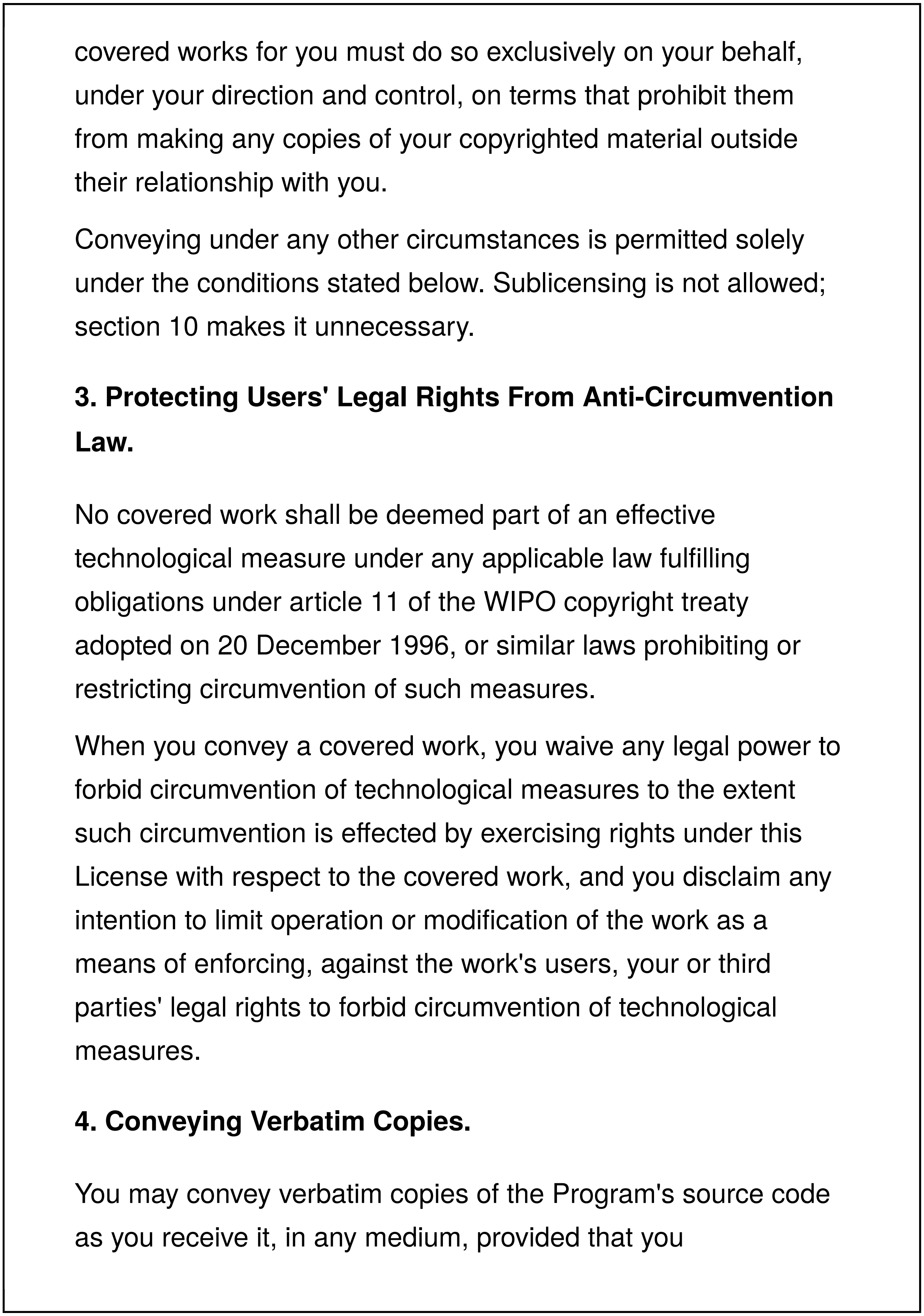
The Corresponding Source need not include anything that users can regenerate automatically from other parts of the Corresponding Source.

The Corresponding Source for a work in source code form is that same work.

2. Basic Permissions.

All rights granted under this License are granted for the term of copyright on the Program, and are irrevocable provided the stated conditions are met. This License explicitly affirms your unlimited permission to run the unmodified Program, subject to section 13. The output from running a covered work is covered by this License only if the output, given its content, constitutes a covered work. This License acknowledges your rights of fair use or other equivalent, as provided by copyright law.

Subject to section 13, you may make, run and propagate covered works that you do not convey, without conditions so long as your license otherwise remains in force. You may convey covered works to others for the sole purpose of having them make modifications exclusively for you, or provide you with facilities for running those works, provided that you comply with the terms of this License in conveying all material for which you do not control copyright. Those thus making or running the



covered works for you must do so exclusively on your behalf, under your direction and control, on terms that prohibit them from making any copies of your copyrighted material outside their relationship with you.

Conveying under any other circumstances is permitted solely under the conditions stated below. Sublicensing is not allowed; section 10 makes it unnecessary.

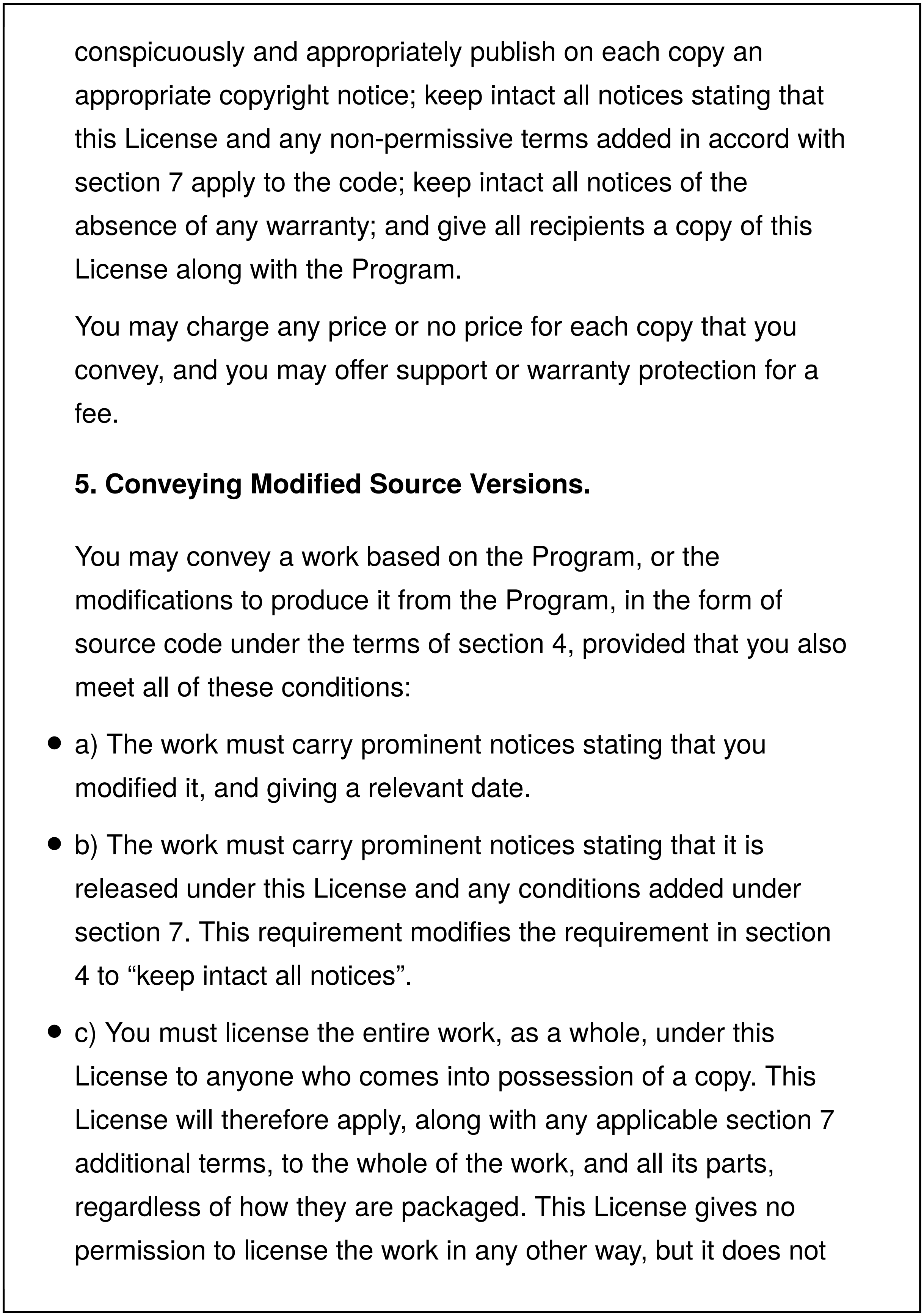
3. Protecting Users' Legal Rights From Anti-Circumvention Law.

No covered work shall be deemed part of an effective technological measure under any applicable law fulfilling obligations under article 11 of the WIPO copyright treaty adopted on 20 December 1996, or similar laws prohibiting or restricting circumvention of such measures.

When you convey a covered work, you waive any legal power to forbid circumvention of technological measures to the extent such circumvention is effected by exercising rights under this License with respect to the covered work, and you disclaim any intention to limit operation or modification of the work as a means of enforcing, against the work's users, your or third parties' legal rights to forbid circumvention of technological measures.

4. Conveying Verbatim Copies.

You may convey verbatim copies of the Program's source code as you receive it, in any medium, provided that you



conspicuously and appropriately publish on each copy an appropriate copyright notice; keep intact all notices stating that this License and any non-permissive terms added in accord with section 7 apply to the code; keep intact all notices of the absence of any warranty; and give all recipients a copy of this License along with the Program.

You may charge any price or no price for each copy that you convey, and you may offer support or warranty protection for a fee.

5. Conveying Modified Source Versions.

You may convey a work based on the Program, or the modifications to produce it from the Program, in the form of source code under the terms of section 4, provided that you also meet all of these conditions:

- a) The work must carry prominent notices stating that you modified it, and giving a relevant date.
- b) The work must carry prominent notices stating that it is released under this License and any conditions added under section 7. This requirement modifies the requirement in section 4 to “keep intact all notices”.
- c) You must license the entire work, as a whole, under this License to anyone who comes into possession of a copy. This License will therefore apply, along with any applicable section 7 additional terms, to the whole of the work, and all its parts, regardless of how they are packaged. This License gives no permission to license the work in any other way, but it does not

invalidate such permission if you have separately received it.

- d) If the work has interactive user interfaces, each must display Appropriate Legal Notices; however, if the Program has interactive interfaces that do not display Appropriate Legal Notices, your work need not make them do so.

A compilation of a covered work with other separate and independent works, which are not by their nature extensions of the covered work, and which are not combined with it such as to form a larger program, in or on a volume of a storage or distribution medium, is called an “aggregate” if the compilation and its resulting copyright are not used to limit the access or legal rights of the compilation's users beyond what the individual works permit. Inclusion of a covered work in an aggregate does not cause this License to apply to the other parts of the aggregate.

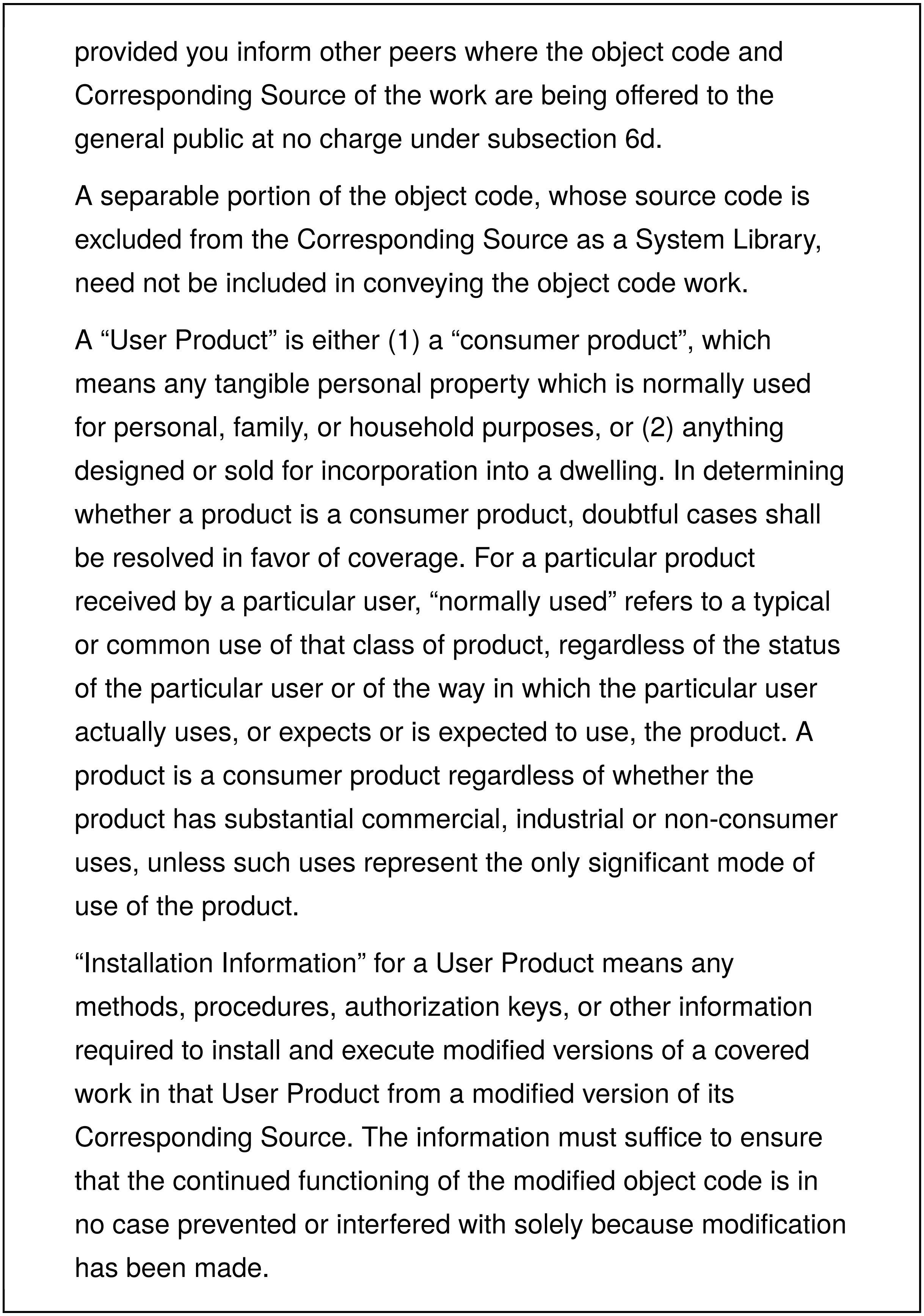
6. Conveying Non-Source Forms.

You may convey a covered work in object code form under the terms of sections 4 and 5, provided that you also convey the machine-readable Corresponding Source under the terms of this License, in one of these ways:

- a) Convey the object code in, or embodied in, a physical product (including a physical distribution medium), accompanied by the Corresponding Source fixed on a durable physical medium customarily used for software interchange.
- b) Convey the object code in, or embodied in, a physical product (including a physical distribution medium), accompanied by a

written offer, valid for at least three years and valid for as long as you offer spare parts or customer support for that product model, to give anyone who possesses the object code either (1) a copy of the Corresponding Source for all the software in the product that is covered by this License, on a durable physical medium customarily used for software interchange, for a price no more than your reasonable cost of physically performing this conveying of source, or (2) access to copy the Corresponding Source from a network server at no charge.

- c) Convey individual copies of the object code with a copy of the written offer to provide the Corresponding Source. This alternative is allowed only occasionally and noncommercially, and only if you received the object code with such an offer, in accord with subsection 6b.
- d) Convey the object code by offering access from a designated place (gratis or for a charge), and offer equivalent access to the Corresponding Source in the same way through the same place at no further charge. You need not require recipients to copy the Corresponding Source along with the object code. If the place to copy the object code is a network server, the Corresponding Source may be on a different server (operated by you or a third party) that supports equivalent copying facilities, provided you maintain clear directions next to the object code saying where to find the Corresponding Source. Regardless of what server hosts the Corresponding Source, you remain obligated to ensure that it is available for as long as needed to satisfy these requirements.
- e) Convey the object code using peer-to-peer transmission,

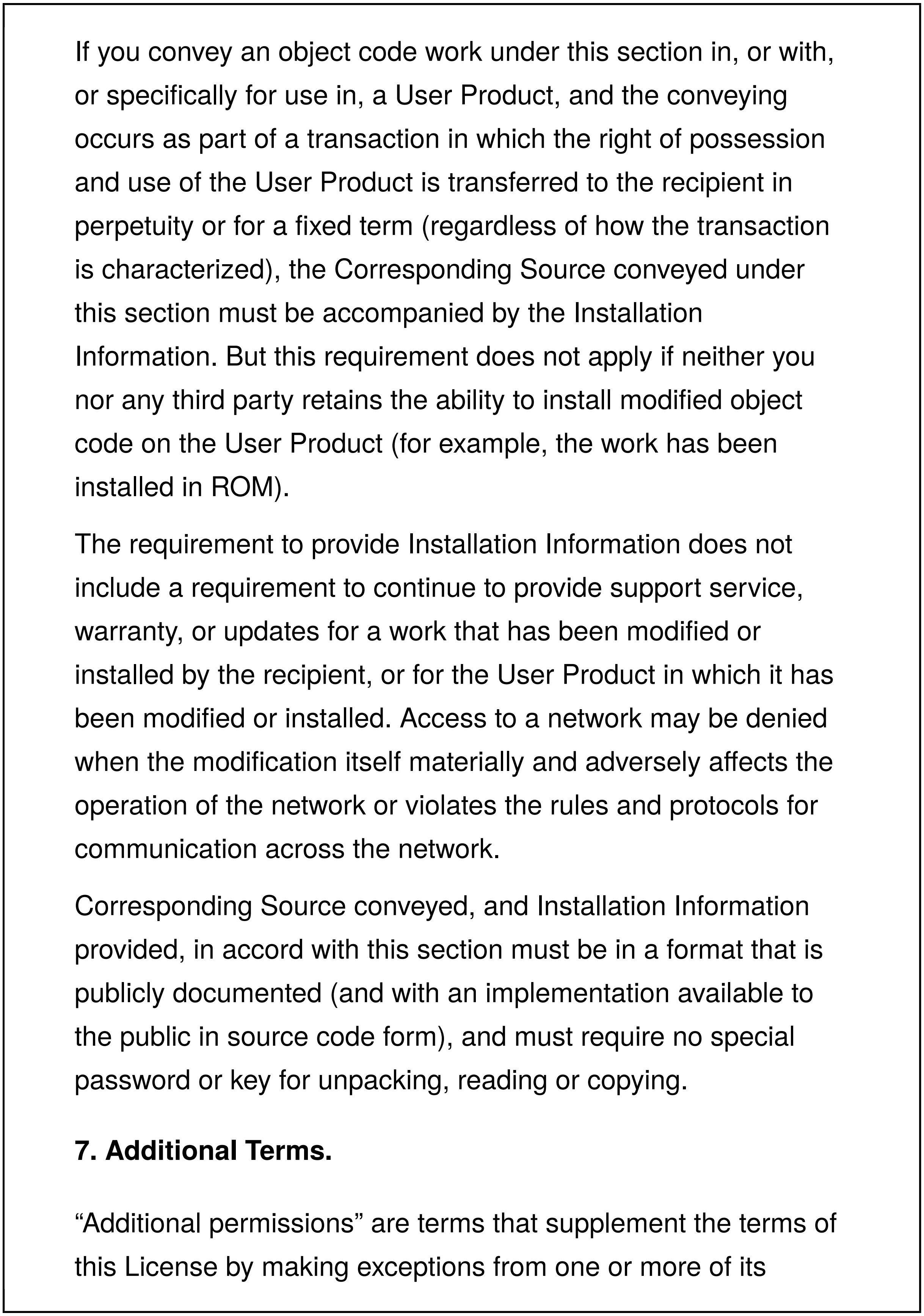


provided you inform other peers where the object code and Corresponding Source of the work are being offered to the general public at no charge under subsection 6d.

A separable portion of the object code, whose source code is excluded from the Corresponding Source as a System Library, need not be included in conveying the object code work.

A “User Product” is either (1) a “consumer product”, which means any tangible personal property which is normally used for personal, family, or household purposes, or (2) anything designed or sold for incorporation into a dwelling. In determining whether a product is a consumer product, doubtful cases shall be resolved in favor of coverage. For a particular product received by a particular user, “normally used” refers to a typical or common use of that class of product, regardless of the status of the particular user or of the way in which the particular user actually uses, or expects or is expected to use, the product. A product is a consumer product regardless of whether the product has substantial commercial, industrial or non-consumer uses, unless such uses represent the only significant mode of use of the product.

“Installation Information” for a User Product means any methods, procedures, authorization keys, or other information required to install and execute modified versions of a covered work in that User Product from a modified version of its Corresponding Source. The information must suffice to ensure that the continued functioning of the modified object code is in no case prevented or interfered with solely because modification has been made.



If you convey an object code work under this section in, or with, or specifically for use in, a User Product, and the conveying occurs as part of a transaction in which the right of possession and use of the User Product is transferred to the recipient in perpetuity or for a fixed term (regardless of how the transaction is characterized), the Corresponding Source conveyed under this section must be accompanied by the Installation Information. But this requirement does not apply if neither you nor any third party retains the ability to install modified object code on the User Product (for example, the work has been installed in ROM).

The requirement to provide Installation Information does not include a requirement to continue to provide support service, warranty, or updates for a work that has been modified or installed by the recipient, or for the User Product in which it has been modified or installed. Access to a network may be denied when the modification itself materially and adversely affects the operation of the network or violates the rules and protocols for communication across the network.

Corresponding Source conveyed, and Installation Information provided, in accord with this section must be in a format that is publicly documented (and with an implementation available to the public in source code form), and must require no special password or key for unpacking, reading or copying.

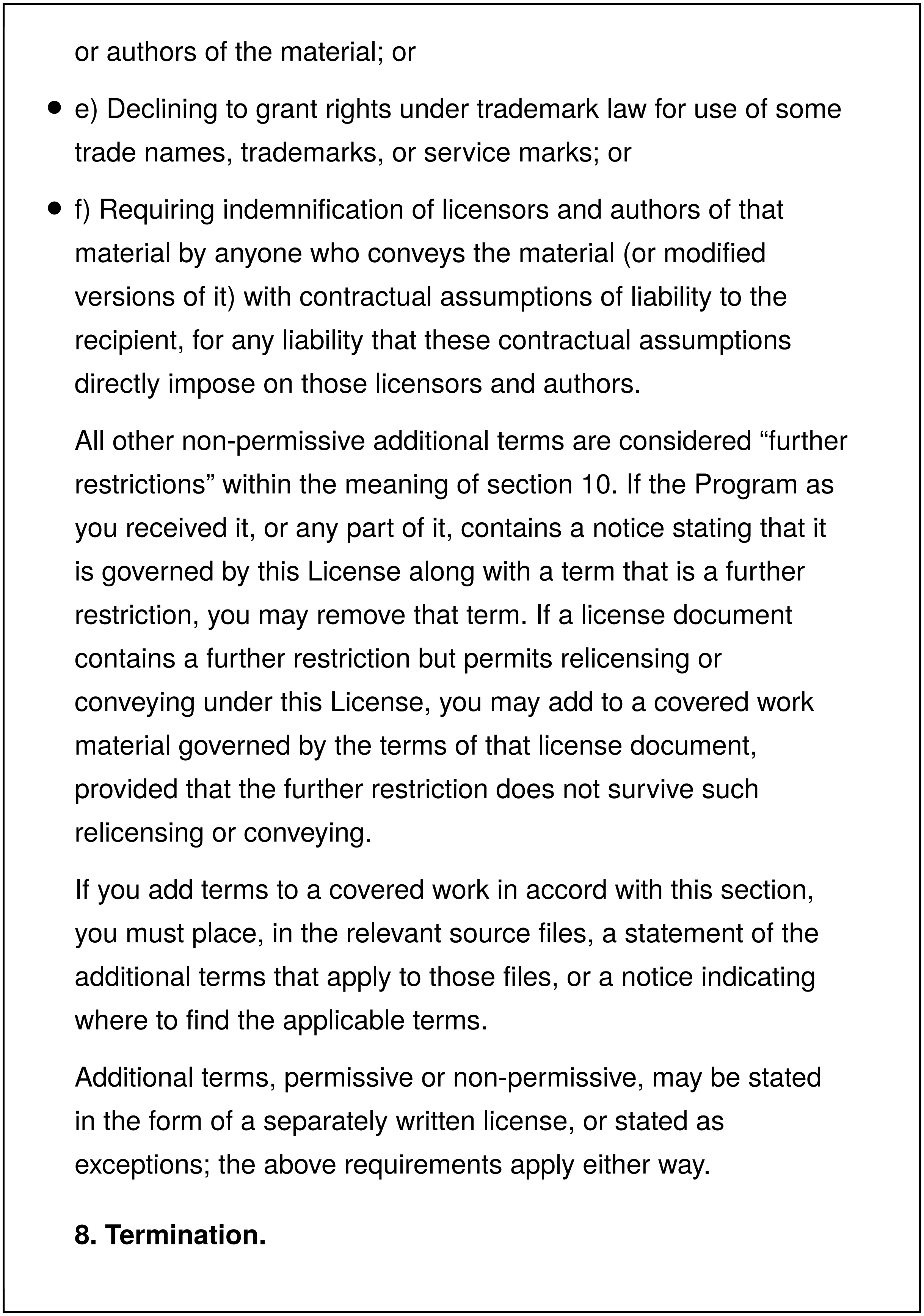
7. Additional Terms.

“Additional permissions” are terms that supplement the terms of this License by making exceptions from one or more of its conditions. Additional permissions that are applicable to the entire Program shall be treated as though they were included in this License, to the extent that they are valid under applicable law. If additional permissions apply only to part of the Program, that part may be used separately under those permissions, but the entire Program remains governed by this License without regard to the additional permissions.

When you convey a copy of a covered work, you may at your option remove any additional permissions from that copy, or from any part of it. (Additional permissions may be written to require their own removal in certain cases when you modify the work.) You may place additional permissions on material, added by you to a covered work, for which you have or can give appropriate copyright permission.

Notwithstanding any other provision of this License, for material you add to a covered work, you may (if authorized by the copyright holders of that material) supplement the terms of this License with terms:

- a) Disclaiming warranty or limiting liability differently from the terms of sections 15 and 16 of this License; or
- b) Requiring preservation of specified reasonable legal notices or author attributions in that material or in the Appropriate Legal Notices displayed by works containing it; or
- c) Prohibiting misrepresentation of the origin of that material, or requiring that modified versions of such material be marked in reasonable ways as different from the original version; or
- d) Limiting the use for publicity purposes of names of licensors



or authors of the material; or

- e) Declining to grant rights under trademark law for use of some trade names, trademarks, or service marks; or
- f) Requiring indemnification of licensors and authors of that material by anyone who conveys the material (or modified versions of it) with contractual assumptions of liability to the recipient, for any liability that these contractual assumptions directly impose on those licensors and authors.

All other non-permissive additional terms are considered “further restrictions” within the meaning of section 10. If the Program as you received it, or any part of it, contains a notice stating that it is governed by this License along with a term that is a further restriction, you may remove that term. If a license document contains a further restriction but permits relicensing or conveying under this License, you may add to a covered work material governed by the terms of that license document, provided that the further restriction does not survive such relicensing or conveying.

If you add terms to a covered work in accord with this section, you must place, in the relevant source files, a statement of the additional terms that apply to those files, or a notice indicating where to find the applicable terms.

Additional terms, permissive or non-permissive, may be stated in the form of a separately written license, or stated as exceptions; the above requirements apply either way.

8. Termination.

You may not propagate or modify a covered work except as expressly provided under this License. Any attempt otherwise to propagate or modify it is void, and will automatically terminate your rights under this License (including any patent licenses granted under the third paragraph of section 11).

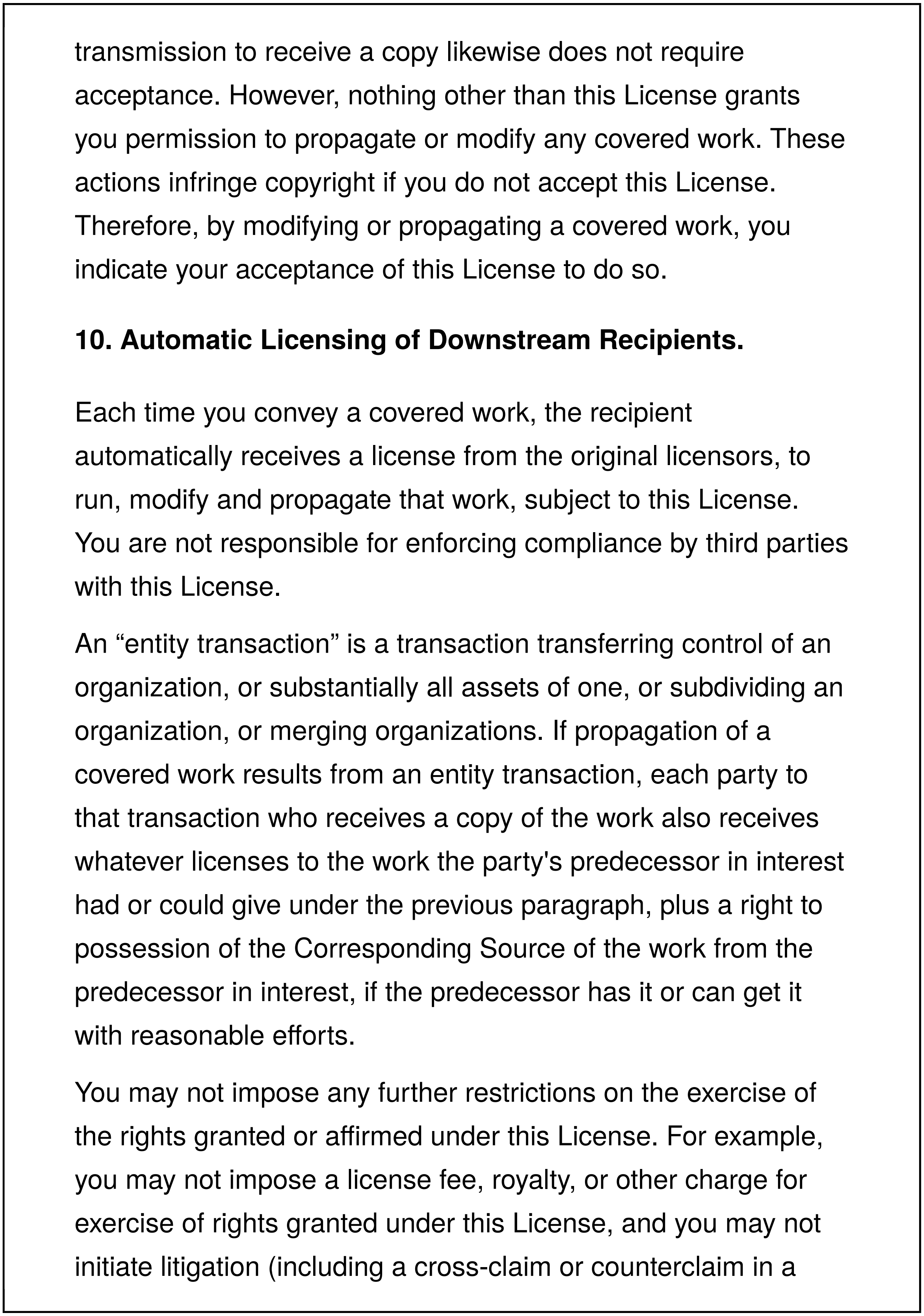
However, if you cease all violation of this License, then your license from a particular copyright holder is reinstated (a) provisionally, unless and until the copyright holder explicitly and finally terminates your license, and (b) permanently, if the copyright holder fails to notify you of the violation by some reasonable means prior to 60 days after the cessation.

Moreover, your license from a particular copyright holder is reinstated permanently if the copyright holder notifies you of the violation by some reasonable means, this is the first time you have received notice of violation of this License (for any work) from that copyright holder, and you cure the violation prior to 30 days after your receipt of the notice.

Termination of your rights under this section does not terminate the licenses of parties who have received copies or rights from you under this License. If your rights have been terminated and not permanently reinstated, you do not qualify to receive new licenses for the same material under section 10.

9. Acceptance Not Required for Having Copies.

You are not required to accept this License in order to receive or run a copy of the Program. Ancillary propagation of a covered work occurring solely as a consequence of using peer-to-peer



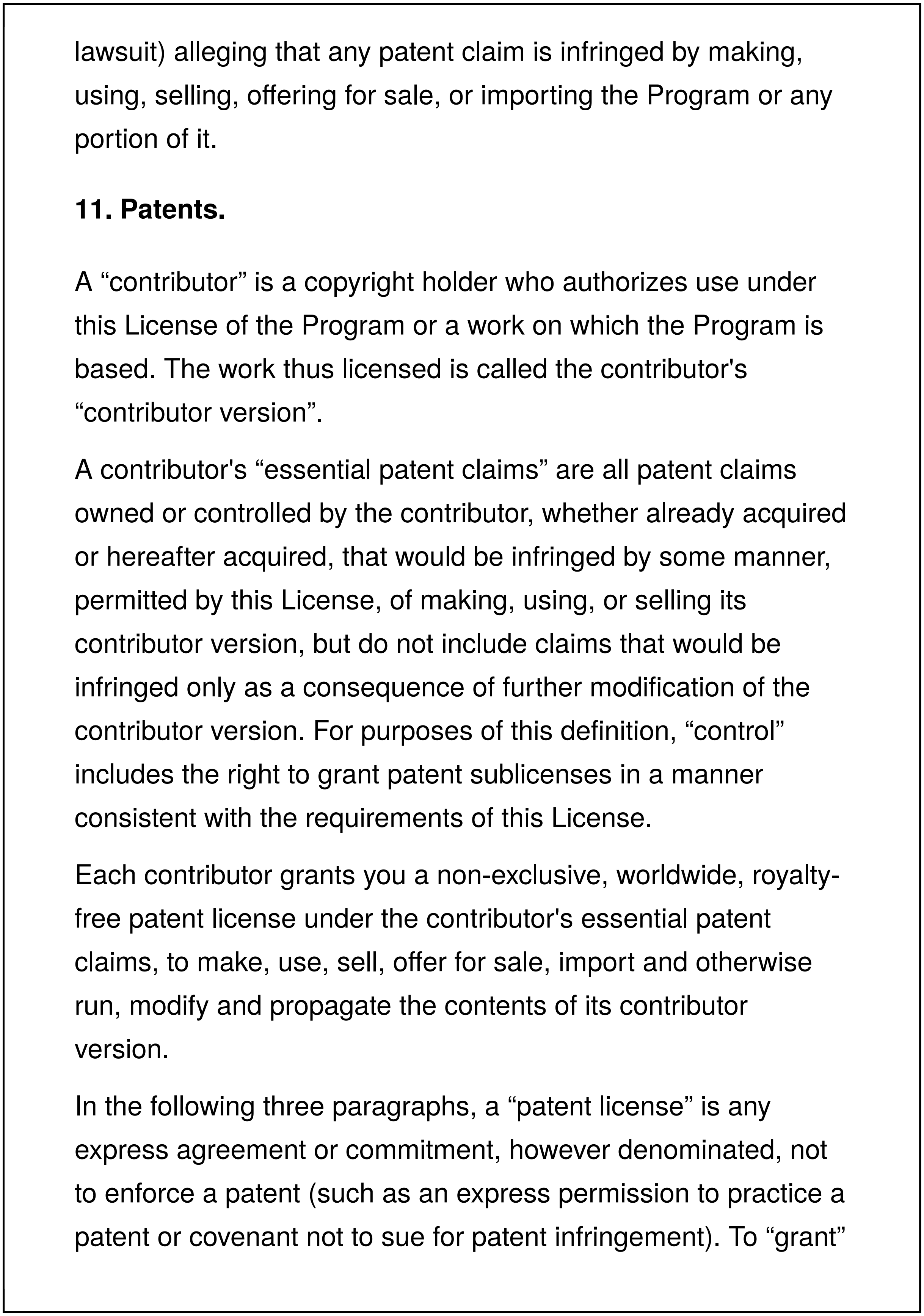
transmission to receive a copy likewise does not require acceptance. However, nothing other than this License grants you permission to propagate or modify any covered work. These actions infringe copyright if you do not accept this License. Therefore, by modifying or propagating a covered work, you indicate your acceptance of this License to do so.

10. Automatic Licensing of Downstream Recipients.

Each time you convey a covered work, the recipient automatically receives a license from the original licensors, to run, modify and propagate that work, subject to this License. You are not responsible for enforcing compliance by third parties with this License.

An “entity transaction” is a transaction transferring control of an organization, or substantially all assets of one, or subdividing an organization, or merging organizations. If propagation of a covered work results from an entity transaction, each party to that transaction who receives a copy of the work also receives whatever licenses to the work the party's predecessor in interest had or could give under the previous paragraph, plus a right to possession of the Corresponding Source of the work from the predecessor in interest, if the predecessor has it or can get it with reasonable efforts.

You may not impose any further restrictions on the exercise of the rights granted or affirmed under this License. For example, you may not impose a license fee, royalty, or other charge for exercise of rights granted under this License, and you may not initiate litigation (including a cross-claim or counterclaim in a



lawsuit) alleging that any patent claim is infringed by making, using, selling, offering for sale, or importing the Program or any portion of it.

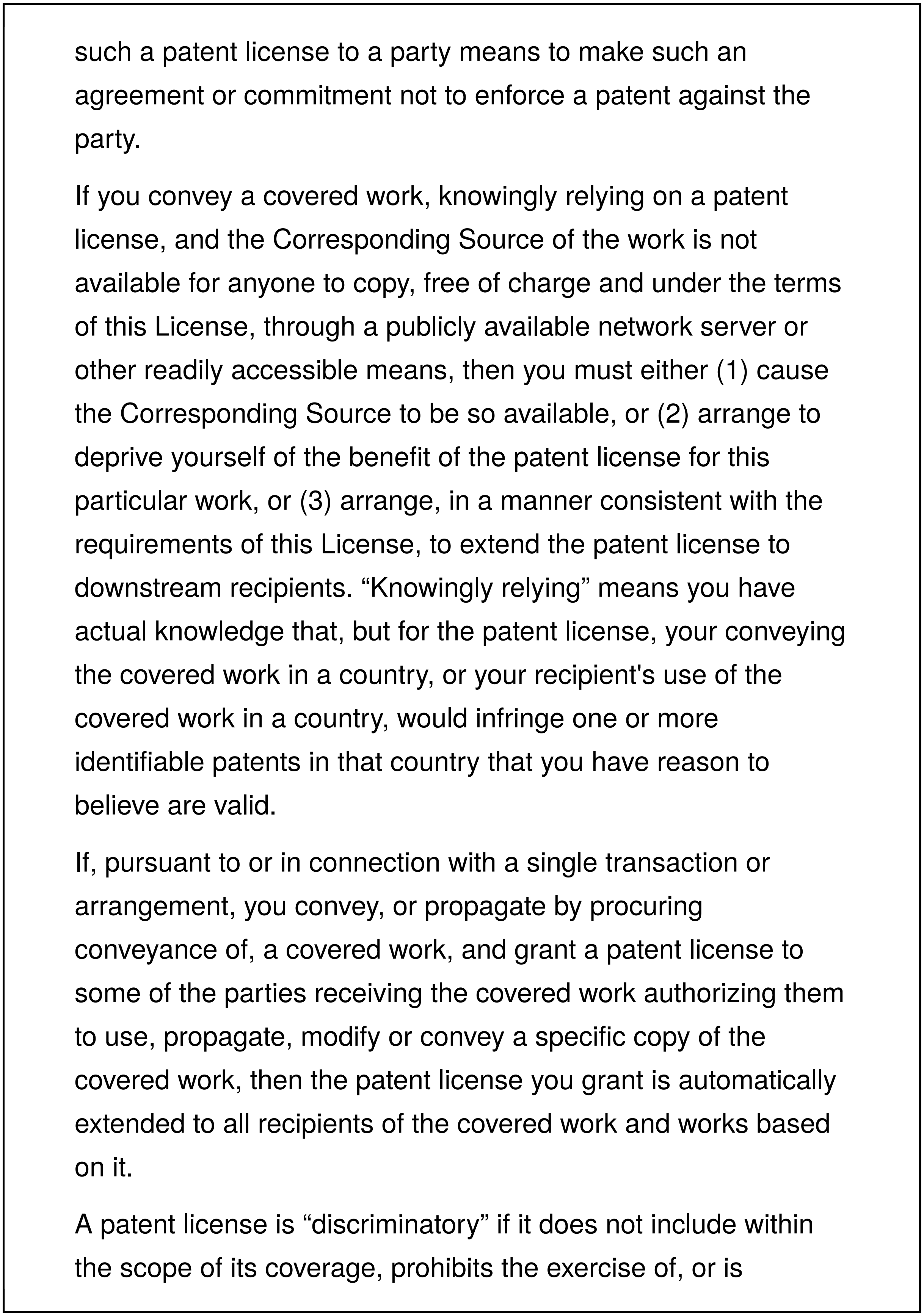
11. Patents.

A “contributor” is a copyright holder who authorizes use under this License of the Program or a work on which the Program is based. The work thus licensed is called the contributor's “contributor version”.

A contributor's “essential patent claims” are all patent claims owned or controlled by the contributor, whether already acquired or hereafter acquired, that would be infringed by some manner, permitted by this License, of making, using, or selling its contributor version, but do not include claims that would be infringed only as a consequence of further modification of the contributor version. For purposes of this definition, “control” includes the right to grant patent sublicenses in a manner consistent with the requirements of this License.

Each contributor grants you a non-exclusive, worldwide, royalty-free patent license under the contributor's essential patent claims, to make, use, sell, offer for sale, import and otherwise run, modify and propagate the contents of its contributor version.

In the following three paragraphs, a “patent license” is any express agreement or commitment, however denominated, not to enforce a patent (such as an express permission to practice a patent or covenant not to sue for patent infringement). To “grant”

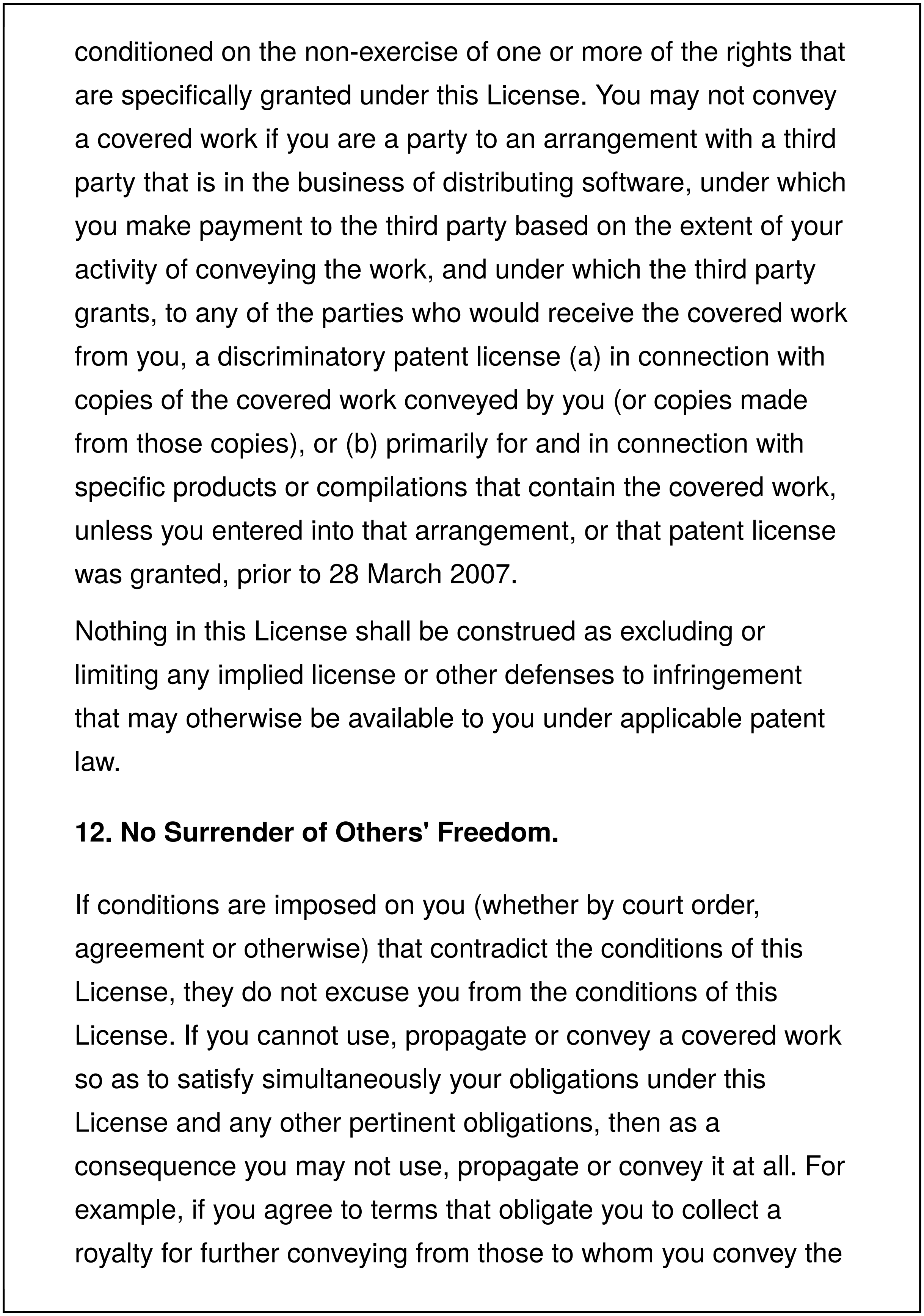


such a patent license to a party means to make such an agreement or commitment not to enforce a patent against the party.

If you convey a covered work, knowingly relying on a patent license, and the Corresponding Source of the work is not available for anyone to copy, free of charge and under the terms of this License, through a publicly available network server or other readily accessible means, then you must either (1) cause the Corresponding Source to be so available, or (2) arrange to deprive yourself of the benefit of the patent license for this particular work, or (3) arrange, in a manner consistent with the requirements of this License, to extend the patent license to downstream recipients. “Knowingly relying” means you have actual knowledge that, but for the patent license, your conveying the covered work in a country, or your recipient's use of the covered work in a country, would infringe one or more identifiable patents in that country that you have reason to believe are valid.

If, pursuant to or in connection with a single transaction or arrangement, you convey, or propagate by procuring conveyance of, a covered work, and grant a patent license to some of the parties receiving the covered work authorizing them to use, propagate, modify or convey a specific copy of the covered work, then the patent license you grant is automatically extended to all recipients of the covered work and works based on it.

A patent license is “discriminatory” if it does not include within the scope of its coverage, prohibits the exercise of, or is

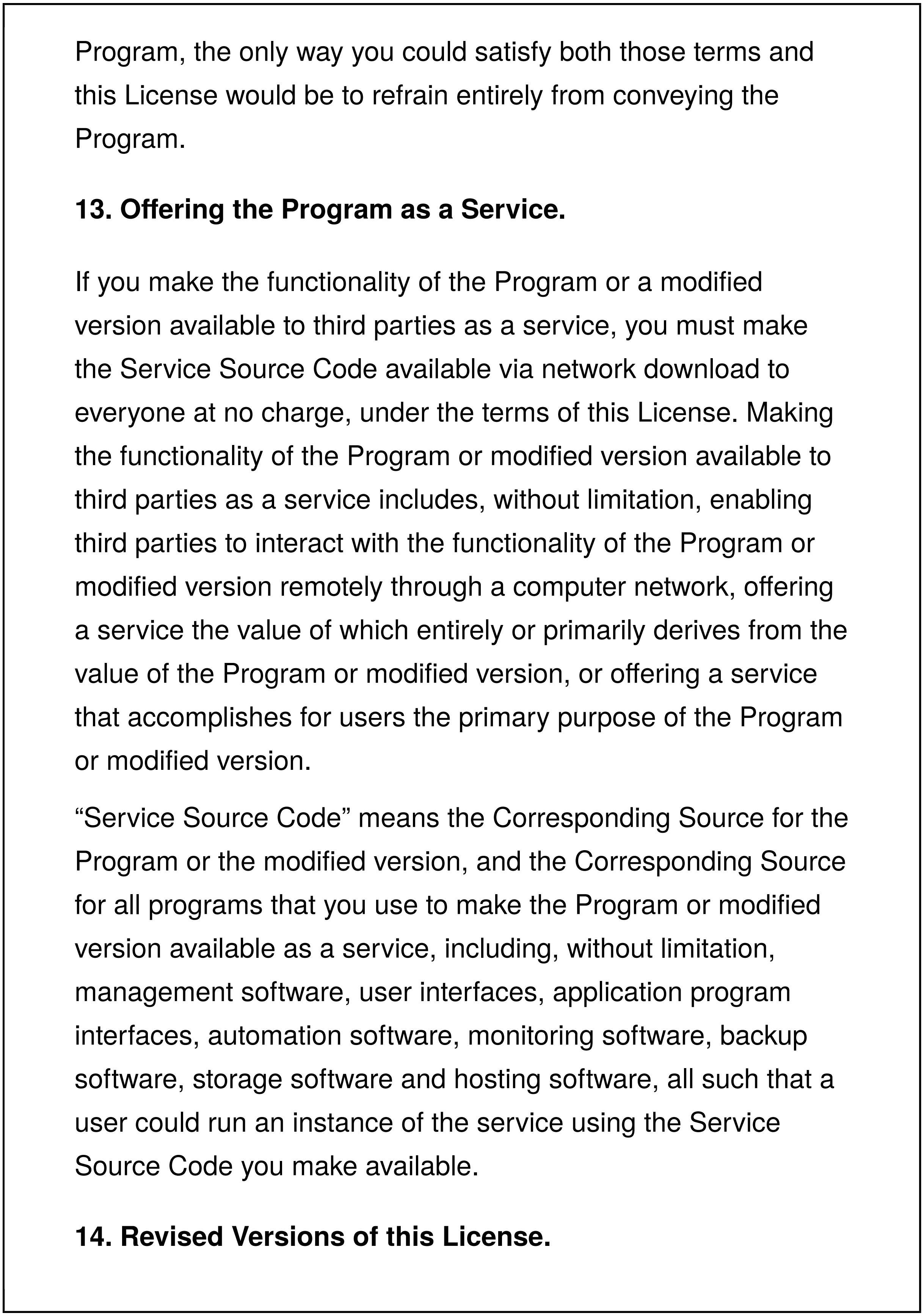


conditioned on the non-exercise of one or more of the rights that are specifically granted under this License. You may not convey a covered work if you are a party to an arrangement with a third party that is in the business of distributing software, under which you make payment to the third party based on the extent of your activity of conveying the work, and under which the third party grants, to any of the parties who would receive the covered work from you, a discriminatory patent license (a) in connection with copies of the covered work conveyed by you (or copies made from those copies), or (b) primarily for and in connection with specific products or compilations that contain the covered work, unless you entered into that arrangement, or that patent license was granted, prior to 28 March 2007.

Nothing in this License shall be construed as excluding or limiting any implied license or other defenses to infringement that may otherwise be available to you under applicable patent law.

12. No Surrender of Others' Freedom.

If conditions are imposed on you (whether by court order, agreement or otherwise) that contradict the conditions of this License, they do not excuse you from the conditions of this License. If you cannot use, propagate or convey a covered work so as to satisfy simultaneously your obligations under this License and any other pertinent obligations, then as a consequence you may not use, propagate or convey it at all. For example, if you agree to terms that obligate you to collect a royalty for further conveying from those to whom you convey the



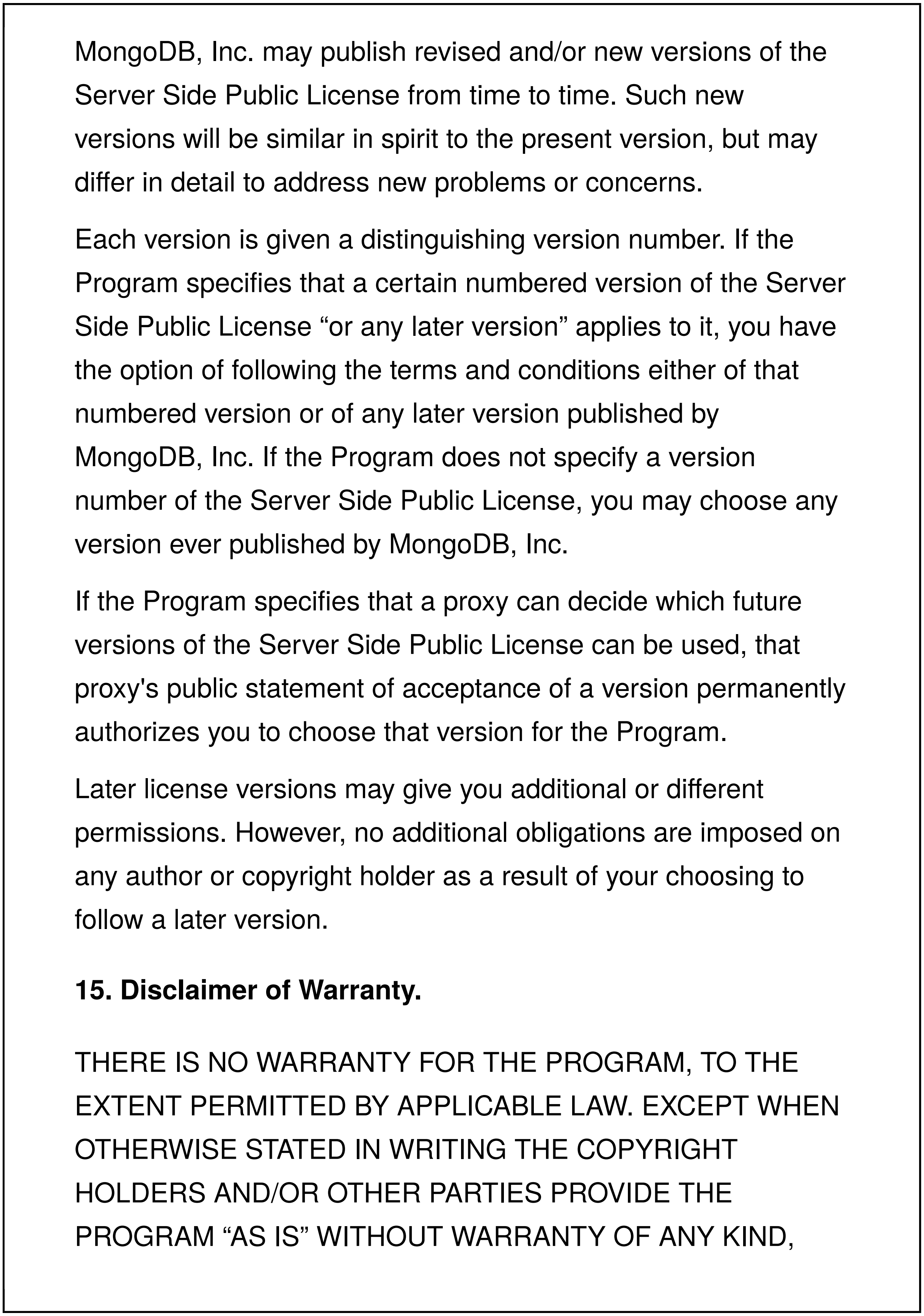
Program, the only way you could satisfy both those terms and this License would be to refrain entirely from conveying the Program.

13. Offering the Program as a Service.

If you make the functionality of the Program or a modified version available to third parties as a service, you must make the Service Source Code available via network download to everyone at no charge, under the terms of this License. Making the functionality of the Program or modified version available to third parties as a service includes, without limitation, enabling third parties to interact with the functionality of the Program or modified version remotely through a computer network, offering a service the value of which entirely or primarily derives from the value of the Program or modified version, or offering a service that accomplishes for users the primary purpose of the Program or modified version.

“Service Source Code” means the Corresponding Source for the Program or the modified version, and the Corresponding Source for all programs that you use to make the Program or modified version available as a service, including, without limitation, management software, user interfaces, application program interfaces, automation software, monitoring software, backup software, storage software and hosting software, all such that a user could run an instance of the service using the Service Source Code you make available.

14. Revised Versions of this License.



MongoDB, Inc. may publish revised and/or new versions of the Server Side Public License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns.

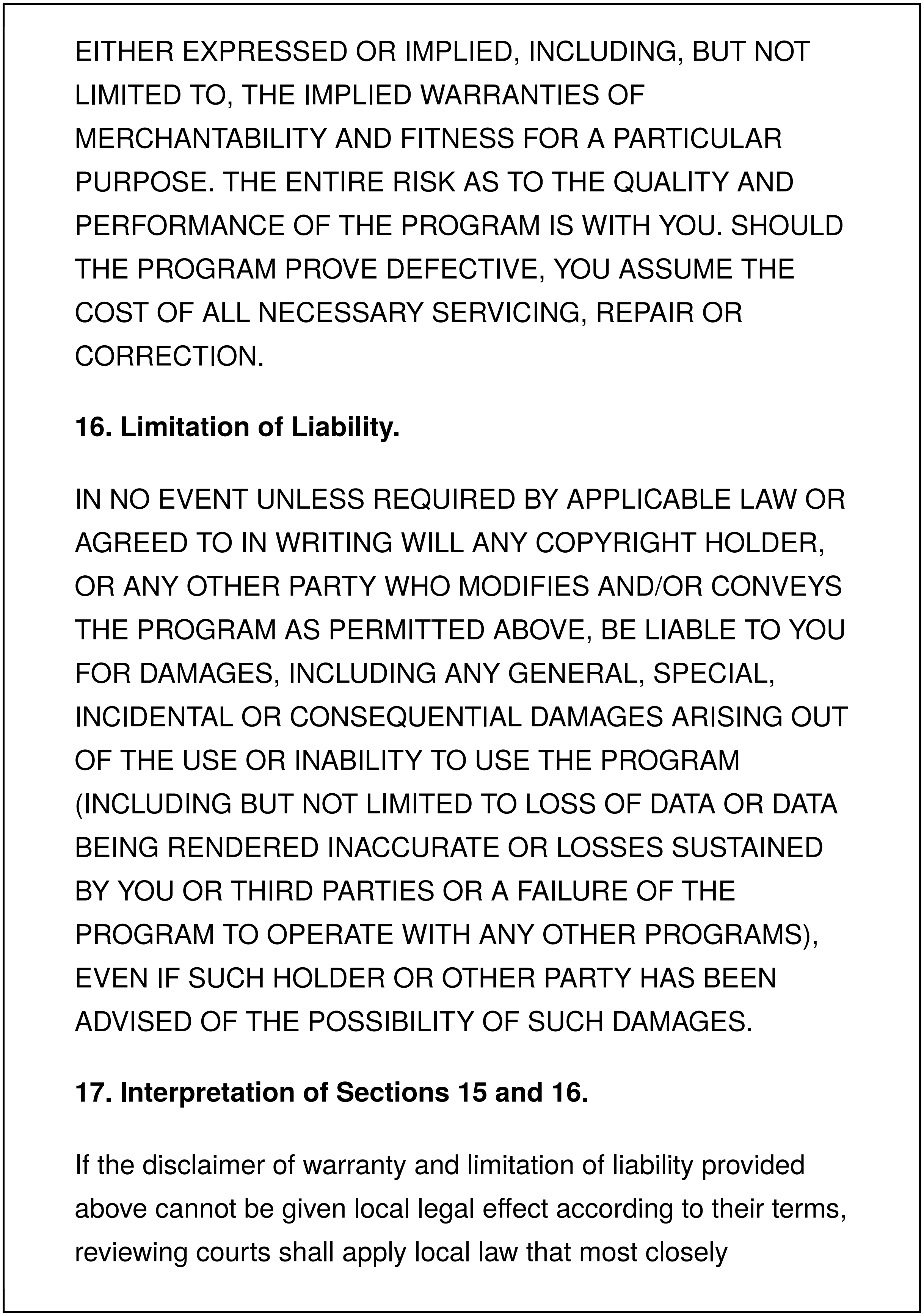
Each version is given a distinguishing version number. If the Program specifies that a certain numbered version of the Server Side Public License “or any later version” applies to it, you have the option of following the terms and conditions either of that numbered version or of any later version published by MongoDB, Inc. If the Program does not specify a version number of the Server Side Public License, you may choose any version ever published by MongoDB, Inc.

If the Program specifies that a proxy can decide which future versions of the Server Side Public License can be used, that proxy's public statement of acceptance of a version permanently authorizes you to choose that version for the Program.

Later license versions may give you additional or different permissions. However, no additional obligations are imposed on any author or copyright holder as a result of your choosing to follow a later version.

15. Disclaimer of Warranty.

THERE IS NO WARRANTY FOR THE PROGRAM, TO THE EXTENT PERMITTED BY APPLICABLE LAW. EXCEPT WHEN OTHERWISE STATED IN WRITING THE COPYRIGHT HOLDERS AND/OR OTHER PARTIES PROVIDE THE PROGRAM “AS IS” WITHOUT WARRANTY OF ANY KIND,



EITHER EXPRESSED OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE. THE ENTIRE RISK AS TO THE QUALITY AND PERFORMANCE OF THE PROGRAM IS WITH YOU. SHOULD THE PROGRAM PROVE DEFECTIVE, YOU ASSUME THE COST OF ALL NECESSARY SERVICING, REPAIR OR CORRECTION.

16. Limitation of Liability.

IN NO EVENT UNLESS REQUIRED BY APPLICABLE LAW OR AGREED TO IN WRITING WILL ANY COPYRIGHT HOLDER, OR ANY OTHER PARTY WHO MODIFIES AND/OR CONVEYS THE PROGRAM AS PERMITTED ABOVE, BE LIABLE TO YOU FOR DAMAGES, INCLUDING ANY GENERAL, SPECIAL, INCIDENTAL OR CONSEQUENTIAL DAMAGES ARISING OUT OF THE USE OR INABILITY TO USE THE PROGRAM (INCLUDING BUT NOT LIMITED TO LOSS OF DATA OR DATA BEING RENDERED INACCURATE OR LOSSES SUSTAINED BY YOU OR THIRD PARTIES OR A FAILURE OF THE PROGRAM TO OPERATE WITH ANY OTHER PROGRAMS), EVEN IF SUCH HOLDER OR OTHER PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

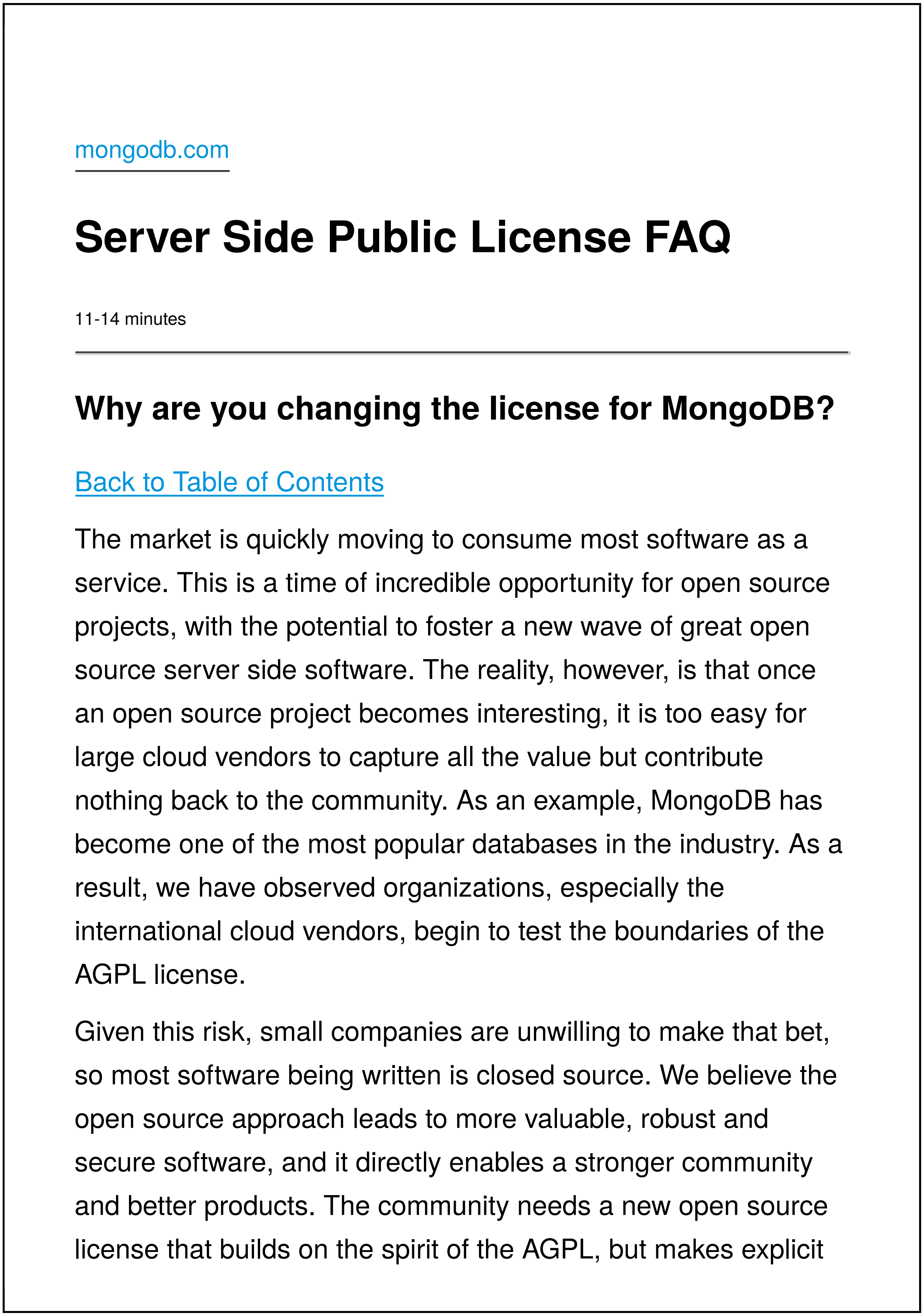
17. Interpretation of Sections 15 and 16.

If the disclaimer of warranty and limitation of liability provided above cannot be given local legal effect according to their terms, reviewing courts shall apply local law that most closely approximates an absolute waiver of all civil liability in connection with the Program, unless a warranty or assumption of liability accompanies a copy of the Program in return for a fee.

END OF TERMS AND CONDITIONS

Chapter 12

Mongo DB Service Side Public License FAQ



mongodb.com

Server Side Public License FAQ

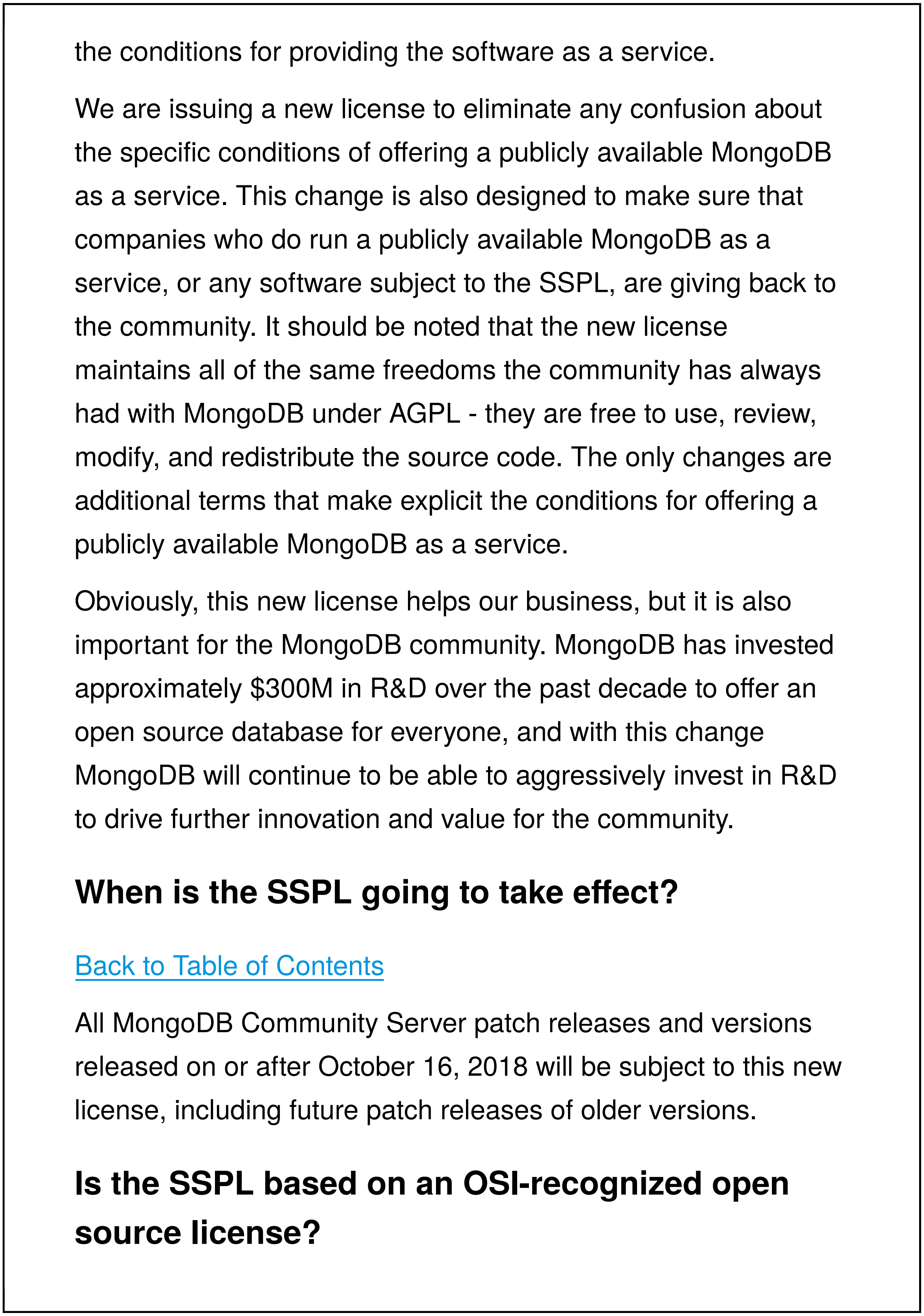
11-14 minutes

Why are you changing the license for MongoDB?

Back to Table of Contents

The market is quickly moving to consume most software as a service. This is a time of incredible opportunity for open source projects, with the potential to foster a new wave of great open source server side software. The reality, however, is that once an open source project becomes interesting, it is too easy for large cloud vendors to capture all the value but contribute nothing back to the community. As an example, MongoDB has become one of the most popular databases in the industry. As a result, we have observed organizations, especially the international cloud vendors, begin to test the boundaries of the AGPL license.

Given this risk, small companies are unwilling to make that bet, so most software being written is closed source. We believe the open source approach leads to more valuable, robust and secure software, and it directly enables a stronger community and better products. The community needs a new open source license that builds on the spirit of the AGPL, but makes explicit



the conditions for providing the software as a service.

We are issuing a new license to eliminate any confusion about the specific conditions of offering a publicly available MongoDB as a service. This change is also designed to make sure that companies who do run a publicly available MongoDB as a service, or any software subject to the SSPL, are giving back to the community. It should be noted that the new license maintains all of the same freedoms the community has always had with MongoDB under AGPL - they are free to use, review, modify, and redistribute the source code. The only changes are additional terms that make explicit the conditions for offering a publicly available MongoDB as a service.

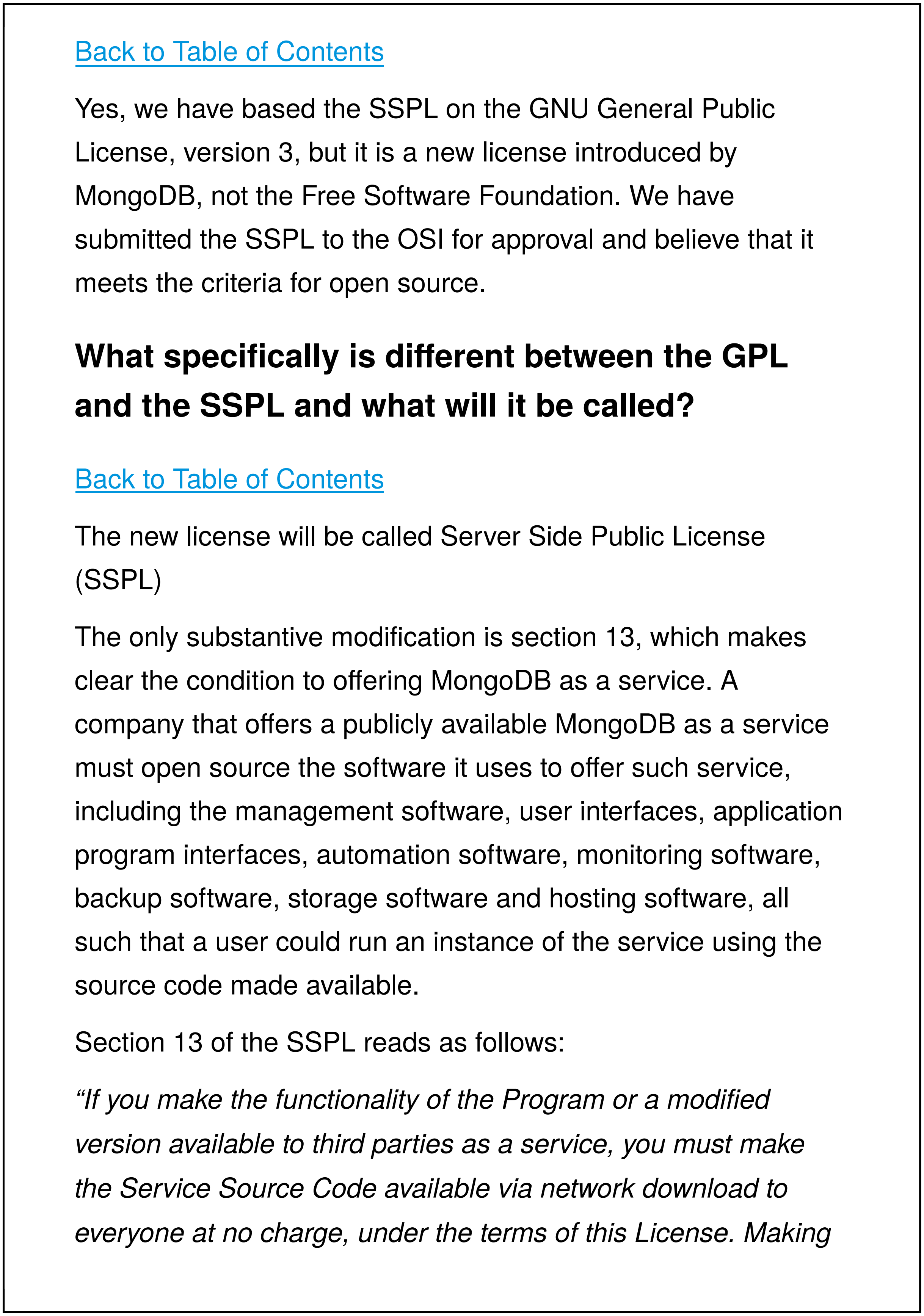
Obviously, this new license helps our business, but it is also important for the MongoDB community. MongoDB has invested approximately $300M in R&D over the past decade to offer an open source database for everyone, and with this change MongoDB will continue to be able to aggressively invest in R&D to drive further innovation and value for the community.

When is the SSPL going to take effect?

Back to Table of Contents

All MongoDB Community Server patch releases and versions released on or after October 16, 2018 will be subject to this new license, including future patch releases of older versions.

Is the SSPL based on an OSI-recognized open source license?



Back to Table of Contents

Yes, we have based the SSPL on the GNU General Public License, version 3, but it is a new license introduced by MongoDB, not the Free Software Foundation. We have submitted the SSPL to the OSI for approval and believe that it meets the criteria for open source.

What specifically is different between the GPL and the SSPL and what will it be called?

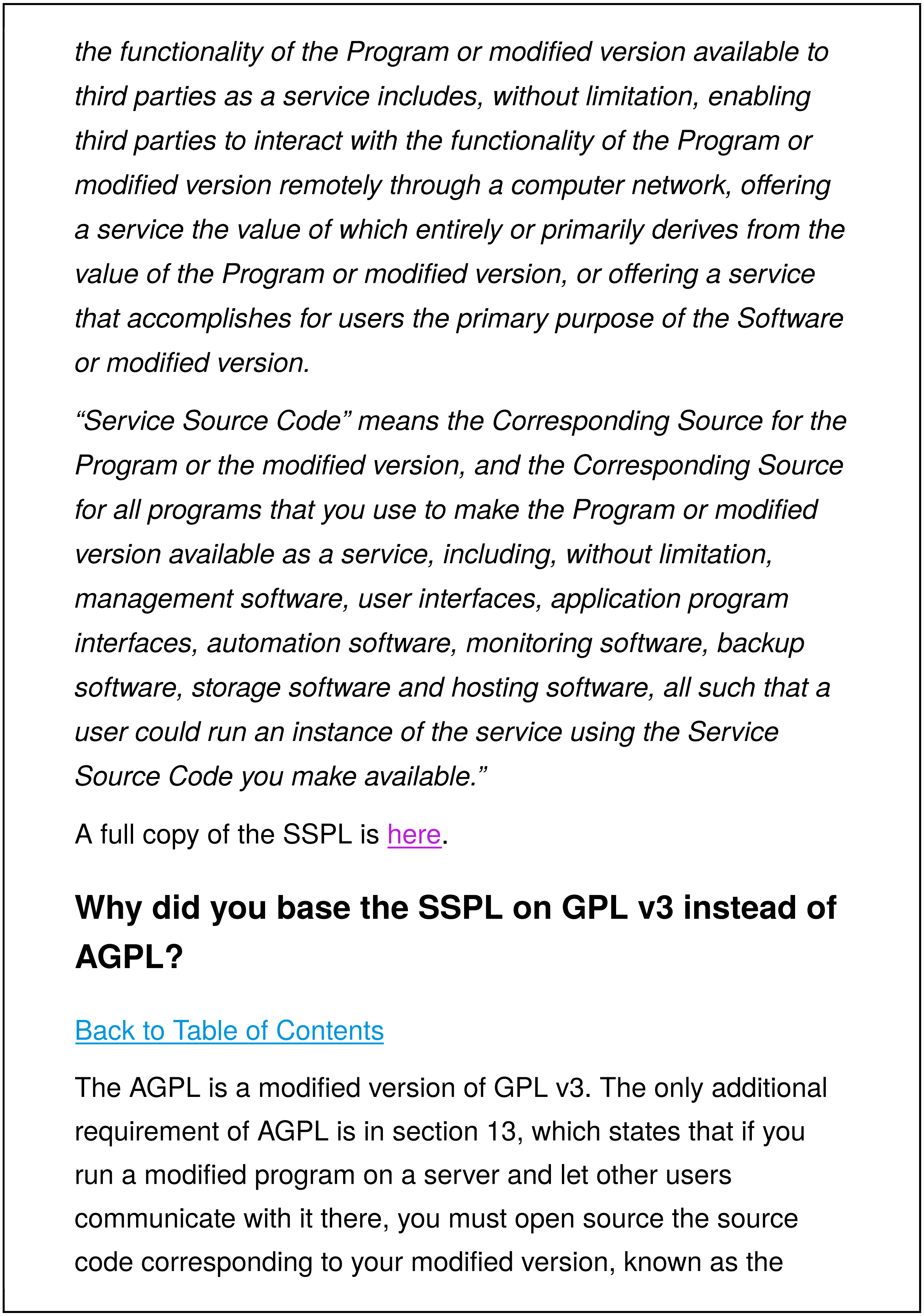
Back to Table of Contents

The new license will be called Server Side Public License (SSPL)

The only substantive modification is section 13, which makes clear the condition to offering MongoDB as a service. A company that offers a publicly available MongoDB as a service must open source the software it uses to offer such service, including the management software, user interfaces, application program interfaces, automation software, monitoring software, backup software, storage software and hosting software, all such that a user could run an instance of the service using the source code made available.

Section 13 of the SSPL reads as follows:

“If you make the functionality of the Program or a modified version available to third parties as a service, you must make the Service Source Code available via network download to everyone at no charge, under the terms of this License. Making



the functionality of the Program or modified version available to third parties as a service includes, without limitation, enabling third parties to interact with the functionality of the Program or modified version remotely through a computer network, offering a service the value of which entirely or primarily derives from the value of the Program or modified version, or offering a service that accomplishes for users the primary purpose of the Software or modified version.

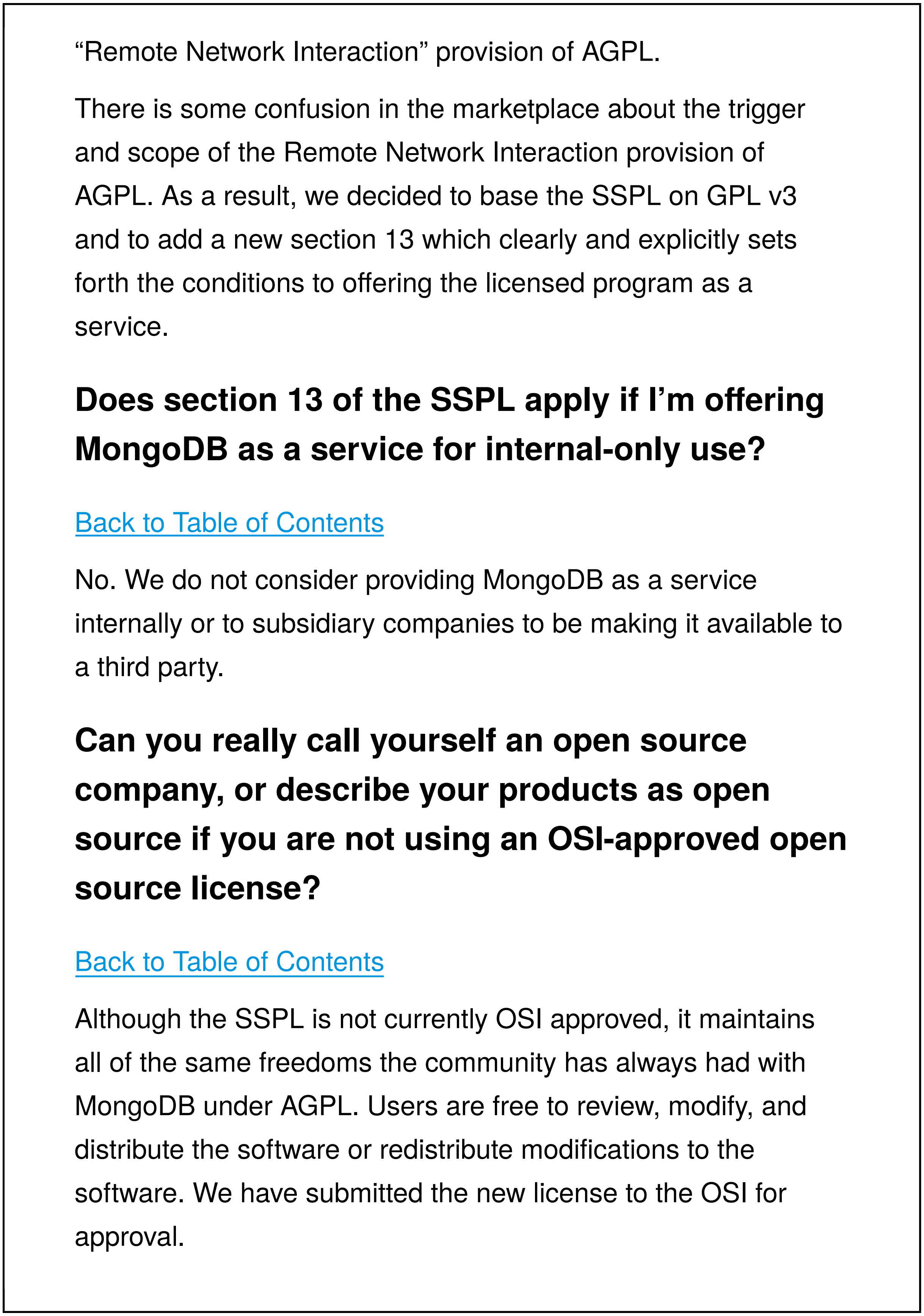
"Service Source Code" means the Corresponding Source for the Program or the modified version, and the Corresponding Source for all programs that you use to make the Program or modified version available as a service, including, without limitation, management software, user interfaces, application program interfaces, automation software, monitoring software, backup software, storage software and hosting software, all such that a user could run an instance of the service using the Service Source Code you make available."

A full copy of the SSPL is here.

Why did you base the SSPL on GPL v3 instead of AGPL?

Back to Table of Contents

The AGPL is a modified version of GPL v3. The only additional requirement of AGPL is in section 13, which states that if you run a modified program on a server and let other users communicate with it there, you must open source the source code corresponding to your modified version, known as the



"Remote Network Interaction" provision of AGPL.

There is some confusion in the marketplace about the trigger and scope of the Remote Network Interaction provision of AGPL. As a result, we decided to base the SSPL on GPL v3 and to add a new section 13 which clearly and explicitly sets forth the conditions to offering the licensed program as a service.

Does section 13 of the SSPL apply if I'm offering MongoDB as a service for internal-only use?

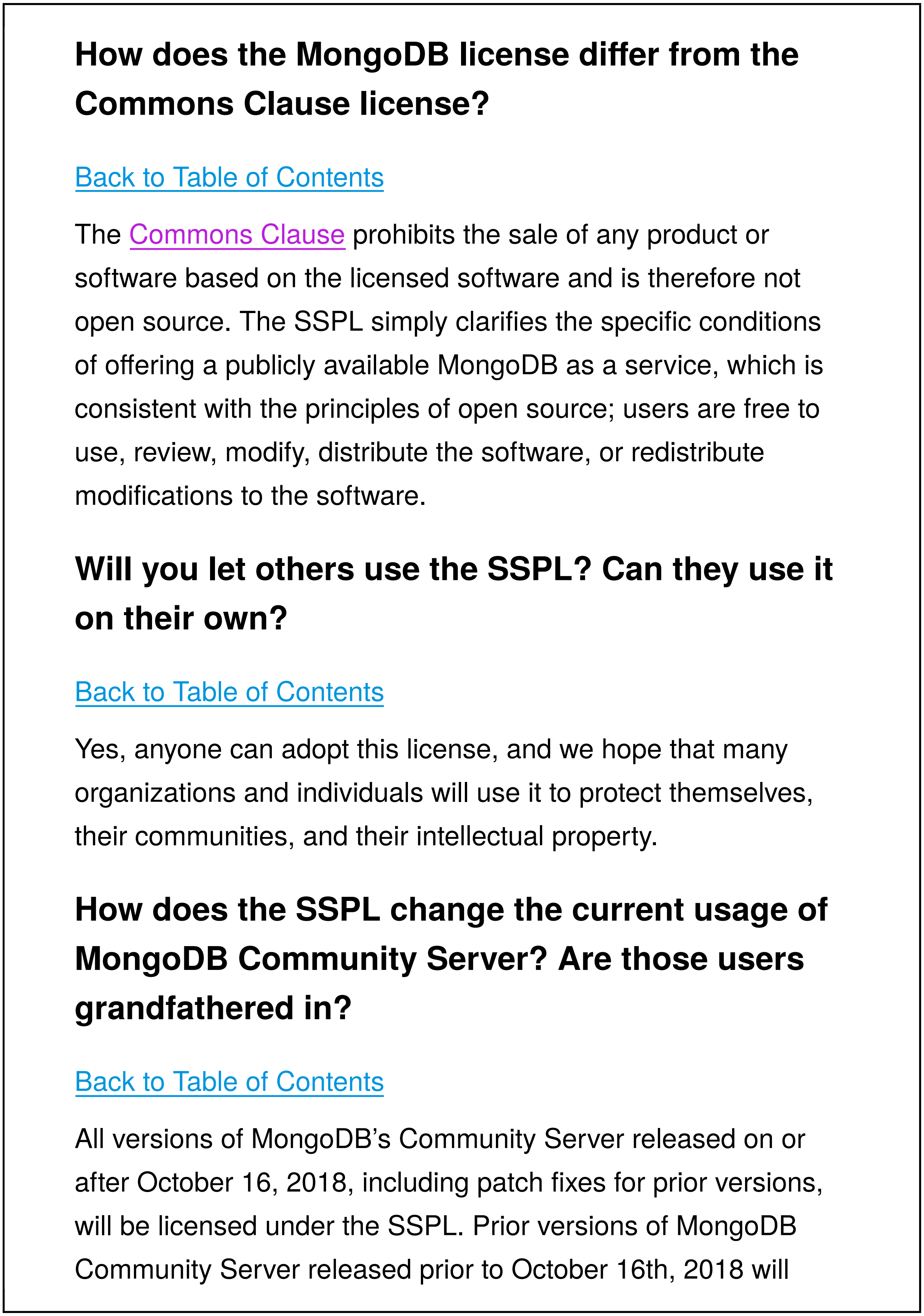
Back to Table of Contents

No. We do not consider providing MongoDB as a service internally or to subsidiary companies to be making it available to a third party.

Can you really call yourself an open source company, or describe your products as open source if you are not using an OSI-approved open source license?

Back to Table of Contents

Although the SSPL is not currently OSI approved, it maintains all of the same freedoms the community has always had with MongoDB under AGPL. Users are free to review, modify, and distribute the software or redistribute modifications to the software. We have submitted the new license to the OSI for approval.



How does the MongoDB license differ from the Commons Clause license?

Back to Table of Contents

The Commons Clause prohibits the sale of any product or software based on the licensed software and is therefore not open source. The SSPL simply clarifies the specific conditions of offering a publicly available MongoDB as a service, which is consistent with the principles of open source; users are free to use, review, modify, distribute the software, or redistribute modifications to the software.

Will you let others use the SSPL? Can they use it on their own?

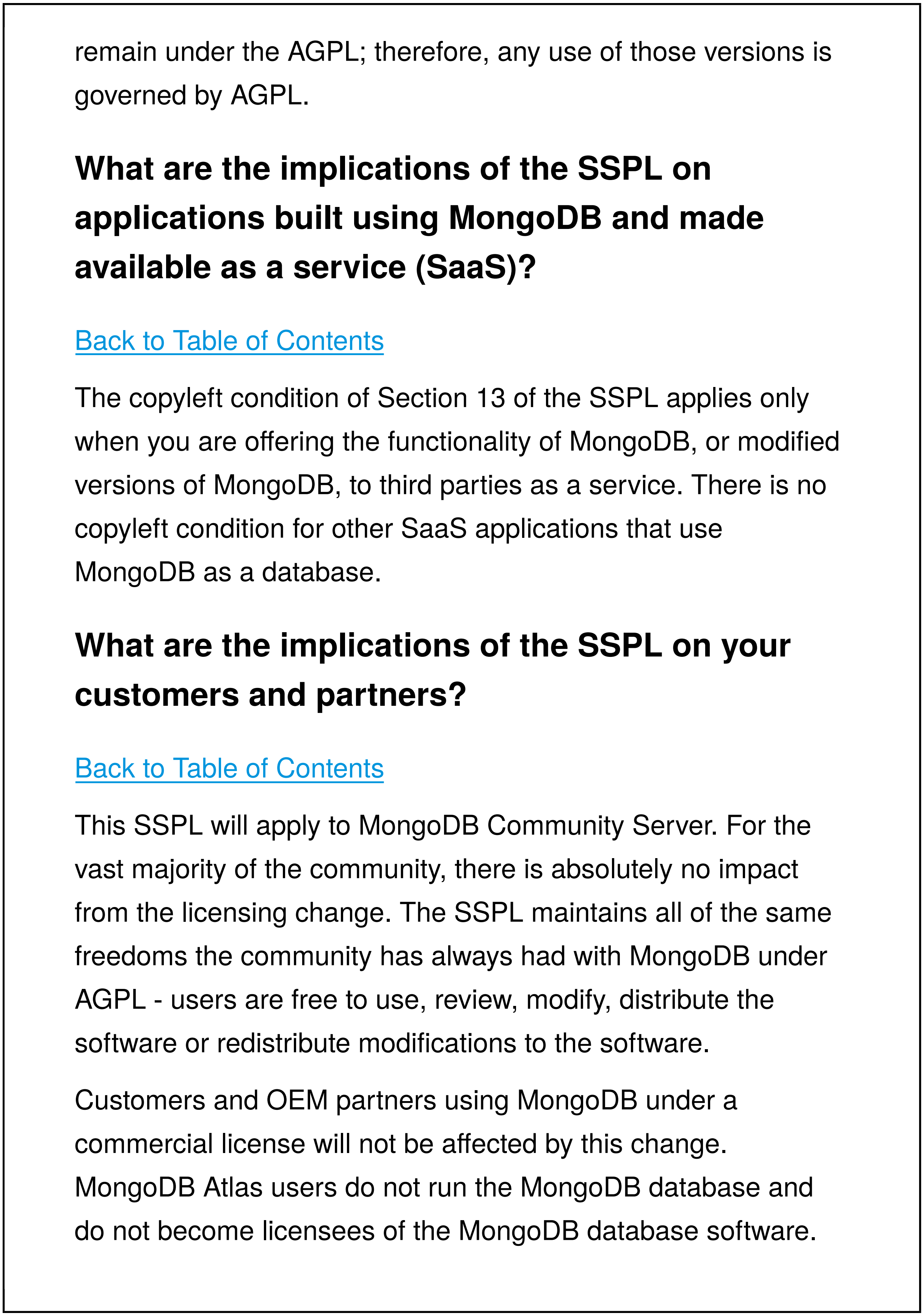
Back to Table of Contents

Yes, anyone can adopt this license, and we hope that many organizations and individuals will use it to protect themselves, their communities, and their intellectual property.

How does the SSPL change the current usage of MongoDB Community Server? Are those users grandfathered in?

Back to Table of Contents

All versions of MongoDB's Community Server released on or after October 16, 2018, including patch fixes for prior versions, will be licensed under the SSPL. Prior versions of MongoDB Community Server released prior to October 16th, 2018 will



remain under the AGPL; therefore, any use of those versions is governed by AGPL.

What are the implications of the SSPL on applications built using MongoDB and made available as a service (SaaS)?

Back to Table of Contents

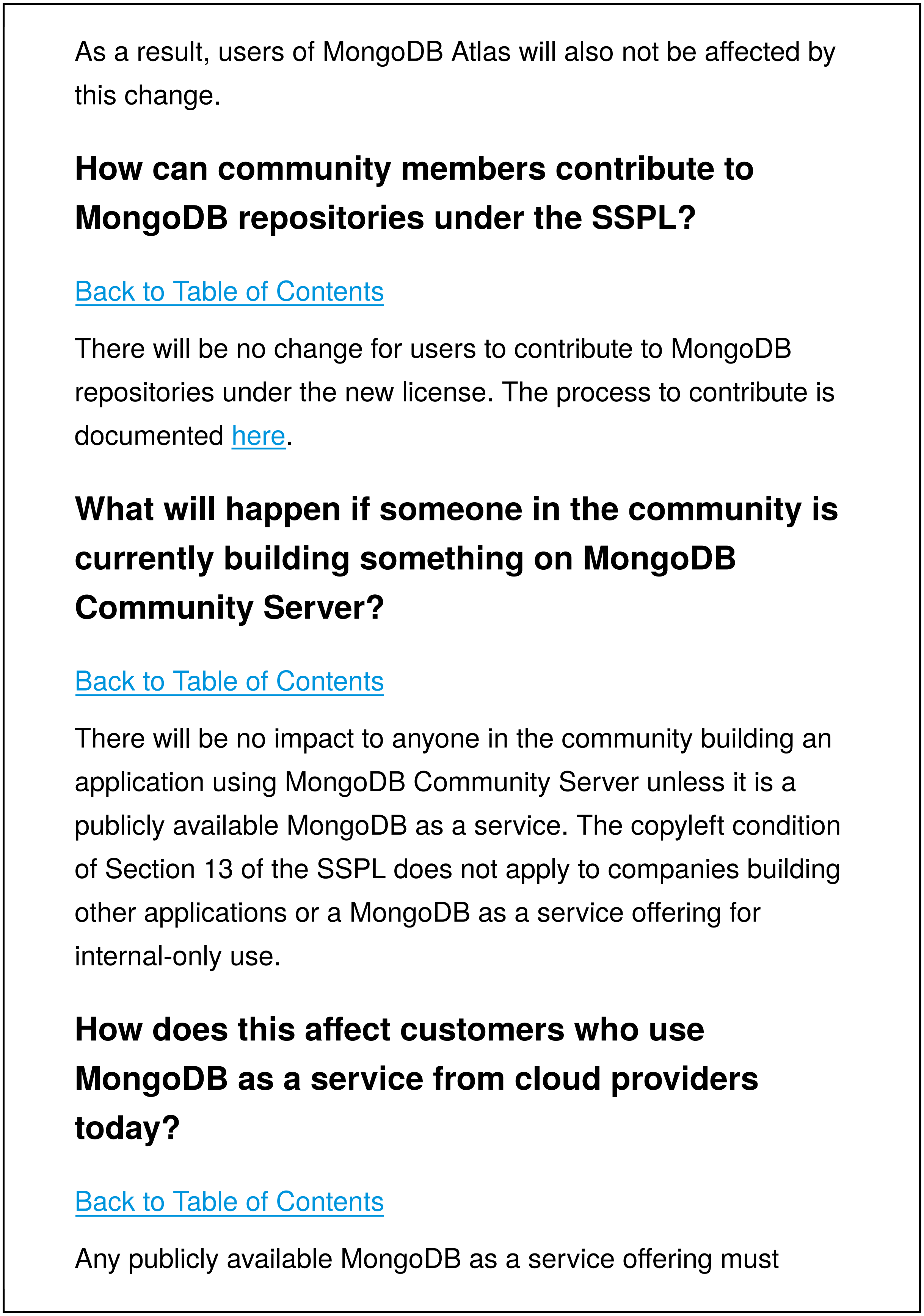
The copyleft condition of Section 13 of the SSPL applies only when you are offering the functionality of MongoDB, or modified versions of MongoDB, to third parties as a service. There is no copyleft condition for other SaaS applications that use MongoDB as a database.

What are the implications of the SSPL on your customers and partners?

Back to Table of Contents

This SSPL will apply to MongoDB Community Server. For the vast majority of the community, there is absolutely no impact from the licensing change. The SSPL maintains all of the same freedoms the community has always had with MongoDB under AGPL - users are free to use, review, modify, distribute the software or redistribute modifications to the software.

Customers and OEM partners using MongoDB under a commercial license will not be affected by this change. MongoDB Atlas users do not run the MongoDB database and do not become licensees of the MongoDB database software.



As a result, users of MongoDB Atlas will also not be affected by this change.

How can community members contribute to MongoDB repositories under the SSPL?

Back to Table of Contents

There will be no change for users to contribute to MongoDB repositories under the new license. The process to contribute is documented here.

What will happen if someone in the community is currently building something on MongoDB Community Server?

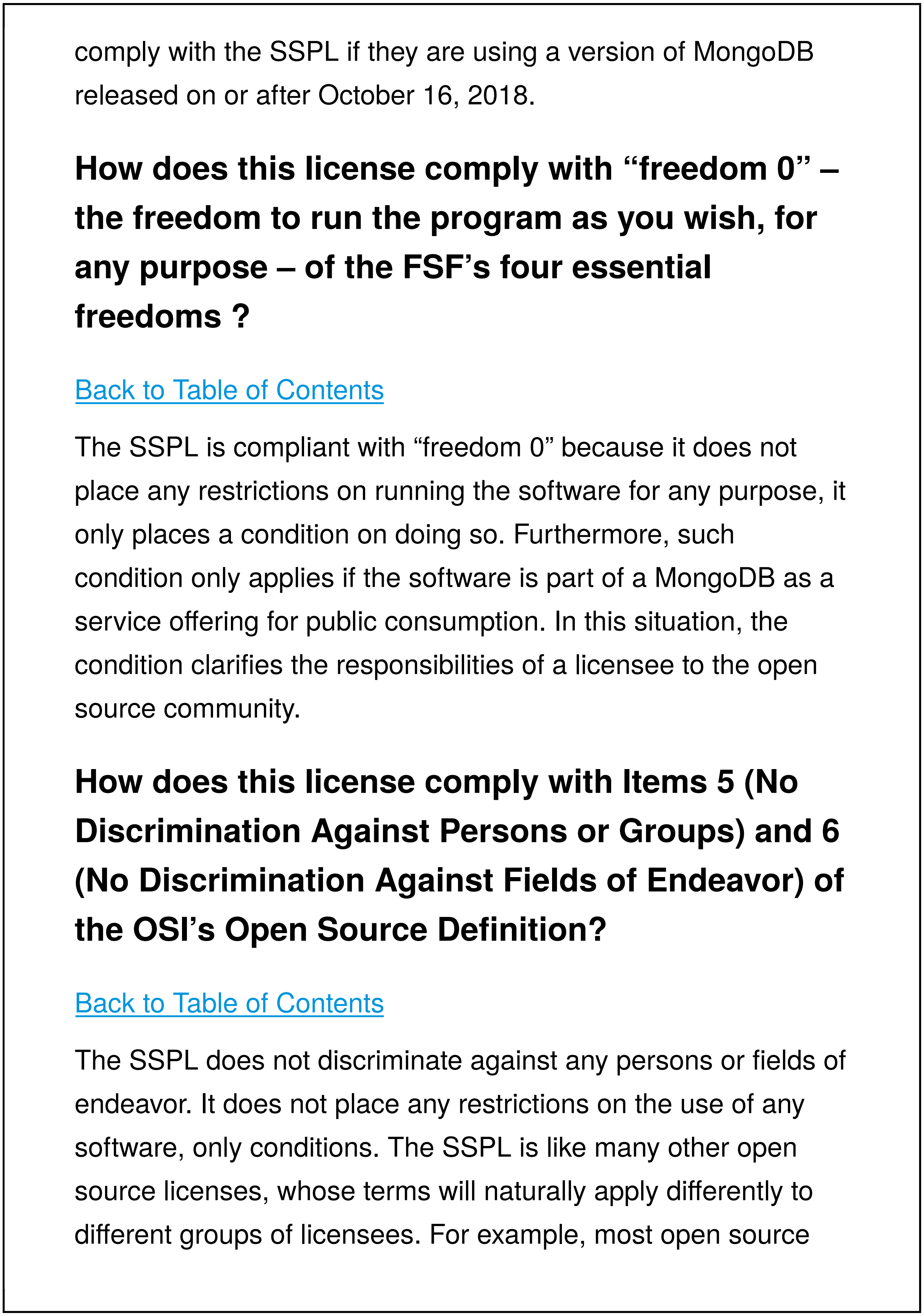
Back to Table of Contents

There will be no impact to anyone in the community building an application using MongoDB Community Server unless it is a publicly available MongoDB as a service. The copyleft condition of Section 13 of the SSPL does not apply to companies building other applications or a MongoDB as a service offering for internal-only use.

How does this affect customers who use MongoDB as a service from cloud providers today?

Back to Table of Contents

Any publicly available MongoDB as a service offering must



comply with the SSPL if they are using a version of MongoDB released on or after October 16, 2018.

How does this license comply with “freedom 0” – the freedom to run the program as you wish, for any purpose – of the FSF’s four essential freedoms ?

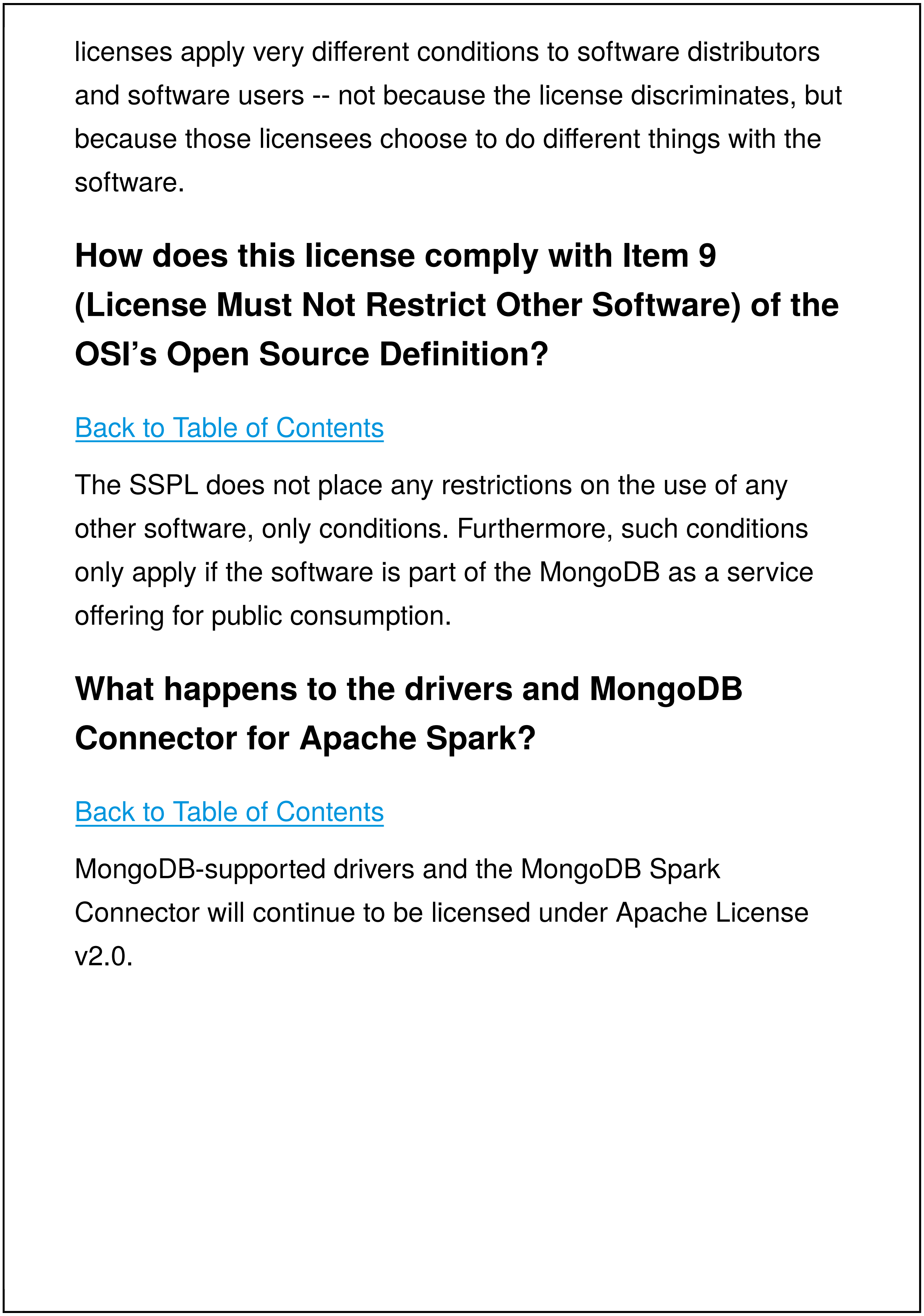
Back to Table of Contents

The SSPL is compliant with “freedom 0” because it does not place any restrictions on running the software for any purpose, it only places a condition on doing so. Furthermore, such condition only applies if the software is part of a MongoDB as a service offering for public consumption. In this situation, the condition clarifies the responsibilities of a licensee to the open source community.

How does this license comply with Items 5 (No Discrimination Against Persons or Groups) and 6 (No Discrimination Against Fields of Endeavor) of the OSI’s Open Source Definition?

Back to Table of Contents

The SSPL does not discriminate against any persons or fields of endeavor. It does not place any restrictions on the use of any software, only conditions. The SSPL is like many other open source licenses, whose terms will naturally apply differently to different groups of licensees. For example, most open source



licenses apply very different conditions to software distributors and software users -- not because the license discriminates, but because those licensees choose to do different things with the software.

How does this license comply with Item 9 (License Must Not Restrict Other Software) of the OSI's Open Source Definition?

Back to Table of Contents

The SSPL does not place any restrictions on the use of any other software, only conditions. Furthermore, such conditions only apply if the software is part of the MongoDB as a service offering for public consumption.

What happens to the drivers and MongoDB Connector for Apache Spark?

Back to Table of Contents

MongoDB-supported drivers and the MongoDB Spark Connector will continue to be licensed under Apache License v2.0.

Chapter 13

Alice v. CLS (SFLC, FSF, OSI Amici Curiae)

No. 13-298

IN THE
Supreme Court of the United States

ALICE CORPORATION PTY. LTD.,
Petitioner,
v.

CLS BANK INTERNATIONAL
AND CLS SERVICES LTD.,
Respondents.

**On Writ of Certiorari to the
United States Court of Appeals
for the Federal Circuit**

BRIEF OF SOFTWARE FREEDOM LAW CENTER, FREE SOFTWARE FOUNDATION, AND OPEN SOURCE INITIATIVE AS *AMICI CURIAE* IN SUPPORT OF RESPONDENTS

EBEN MOGLEN
Counsel of record
MISHI CHOUDHARY
JONATHAN D. BEAN
Software Freedom Law Center
1995 Broadway, 17th Floor
New York, New York 10023
moglen@softwarefreedom.org
212-461-1900

February 27, 2014

QUESTION PRESENTED

Whether claims to computer-implemented inventions—including claims to systems and machines, processes, and items of manufacture—are directed to patent-eligible subject matter within the meaning of 35 U.S.C. § 101 as interpreted by this Court?

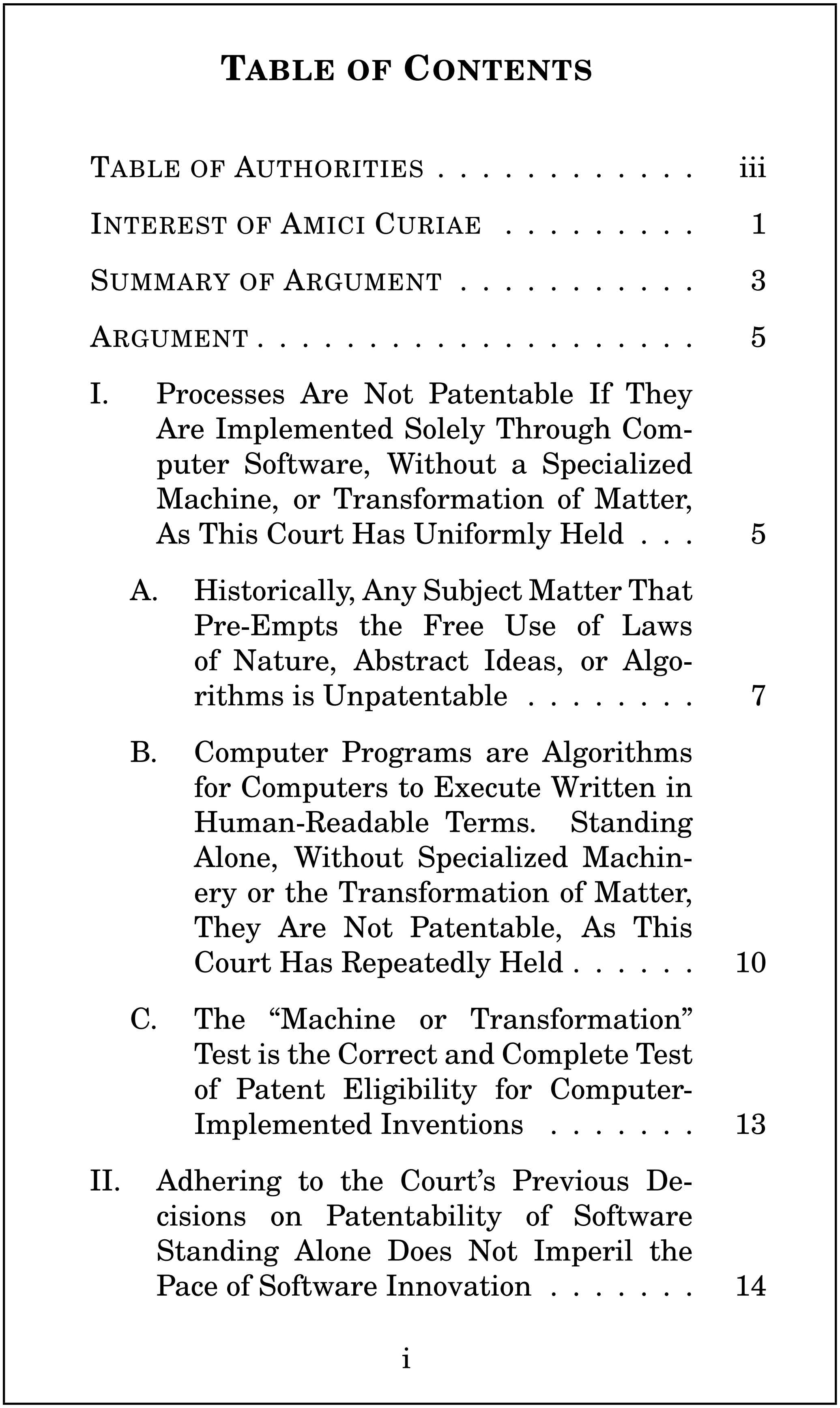


TABLE OF CONTENTS

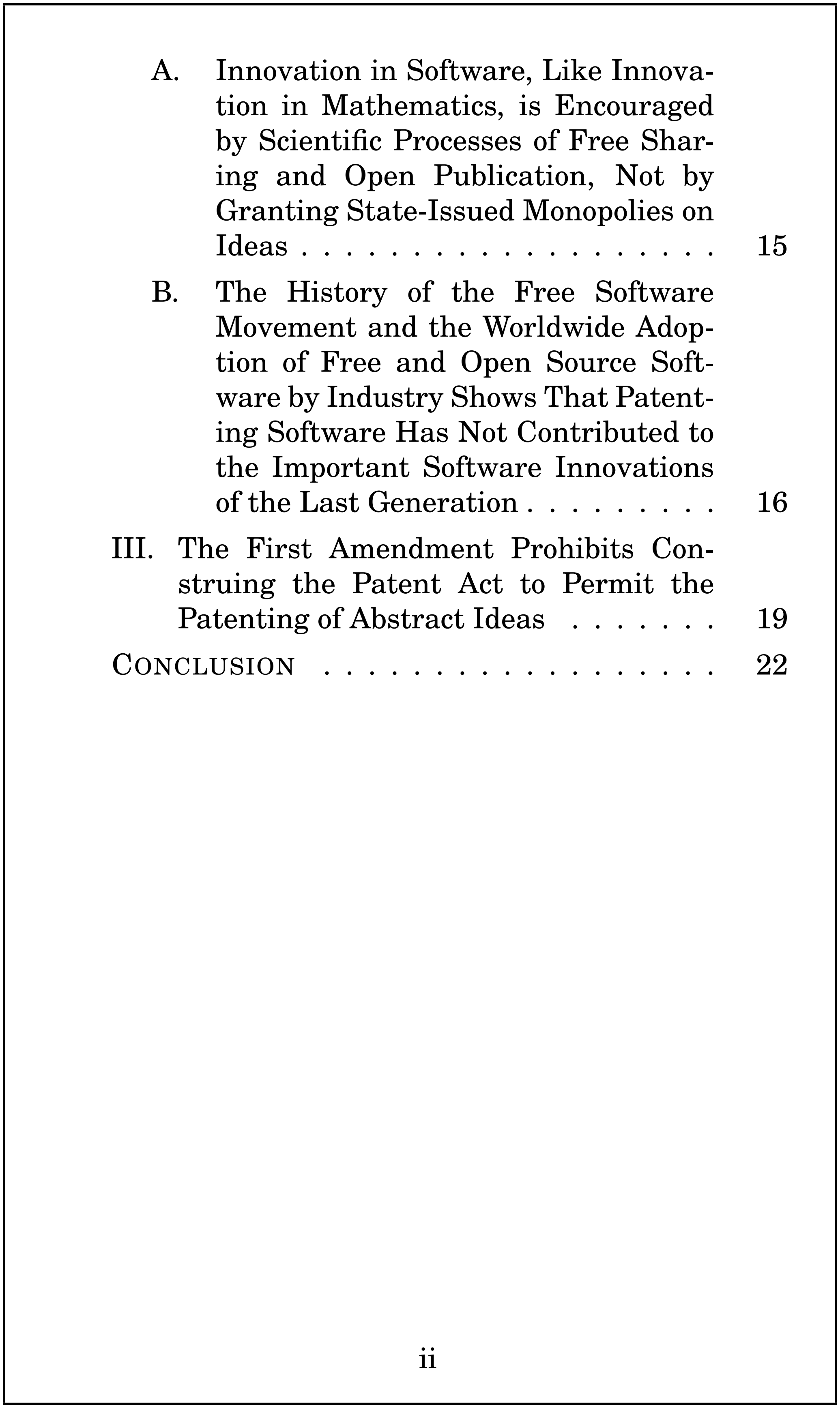
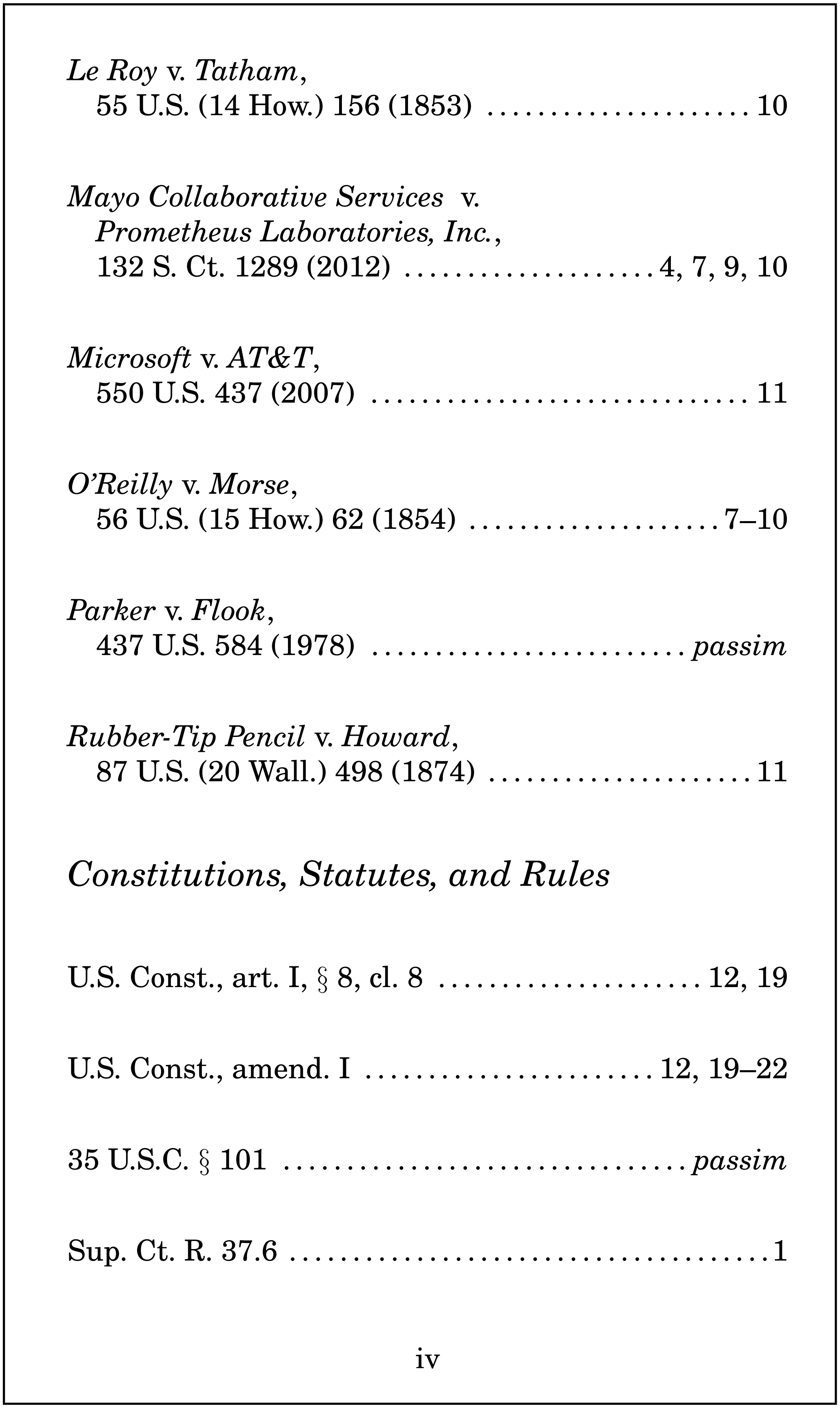


Table of Authorities

Cases



Le Roy v. *Tatham*,
55 U.S. (14 How.) 156 (1853) 10

Mayo Collaborative Services v.
Prometheus Laboratories, Inc.,
132 S. Ct. 1289 (2012) 4, 7, 9, 10

Microsoft v. *AT&T*,
550 U.S. 437 (2007) 11

O'Reilly v. *Morse*,
56 U.S. (15 How.) 62 (1854) 7–10

Parker v. *Flook*,
437 U.S. 584 (1978) .... *passim*

Rubber-Tip Pencil v. *Howard*,
87 U.S. (20 Wall.) 498 (1874) 11

Constitutions, Statutes, and Rules

U.S. Const., art. I, § 8, cl. 8 .... 12, 19

U.S. Const., amend. I 12, 19–22

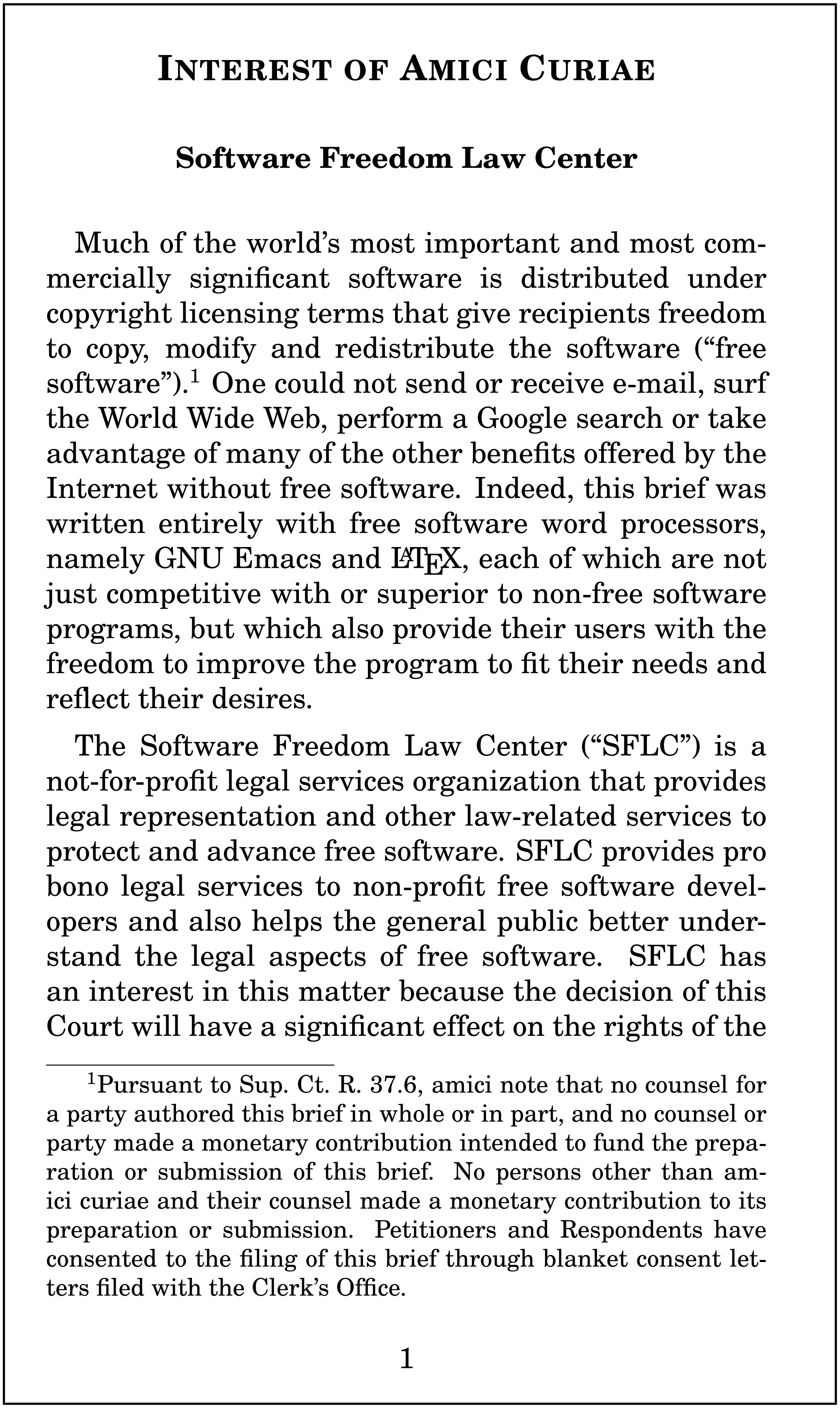
35 U.S.C. § 101 .... *passim*

Sup. Ct. R. 37.6 1

Other Authorities

Brian W. Kernighan and Dennis M. Ritchie, *The C Programming Language* (Prentice Hall 1978) 21, 22

"To Promote the Progress of . . . Useful Arts," Report of the President's Commission on the Patent System (1966) 18, 19



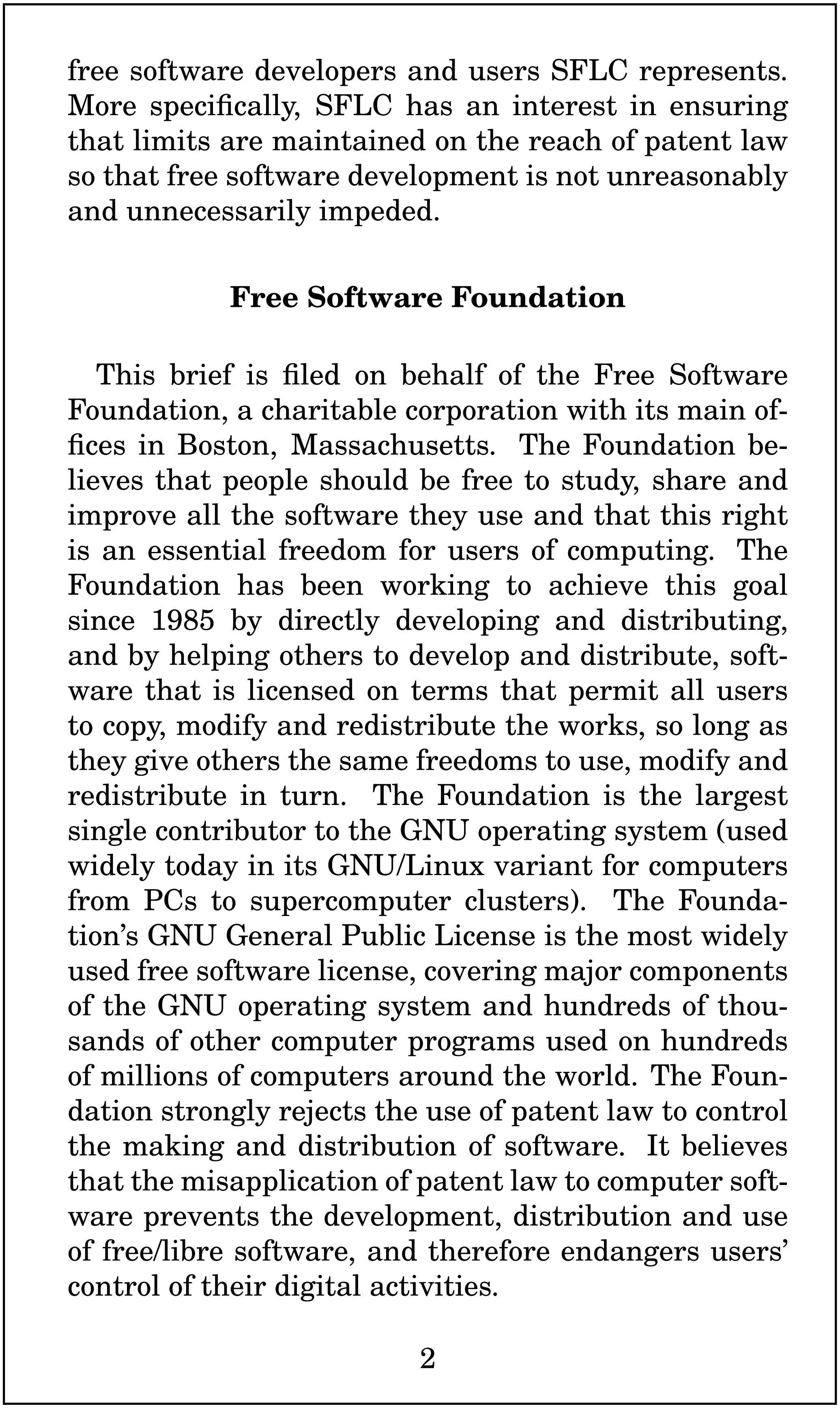
Interest of Amici Curiae

Software Freedom Law Center

Much of the world's most important and most commercially significant software is distributed under copyright licensing terms that give recipients freedom to copy, modify and redistribute the software ("free software").[1] One could not send or receive e-mail, surf the World Wide Web, perform a Google search or take advantage of many of the other benefits offered by the Internet without free software. Indeed, this brief was written entirely with free software word processors, namely GNU Emacs and LaTeX, each of which are not just competitive with or superior to non-free software programs, but which also provide their users with the freedom to improve the program to fit their needs and reflect their desires.

The Software Freedom Law Center ("SFLC") is a not-for-profit legal services organization that provides legal representation and other law-related services to protect and advance free software. SFLC provides pro bono legal services to non-profit free software developers and also helps the general public better understand the legal aspects of free software. SFLC has an interest in this matter because the decision of this Court will have a significant effect on the rights of the

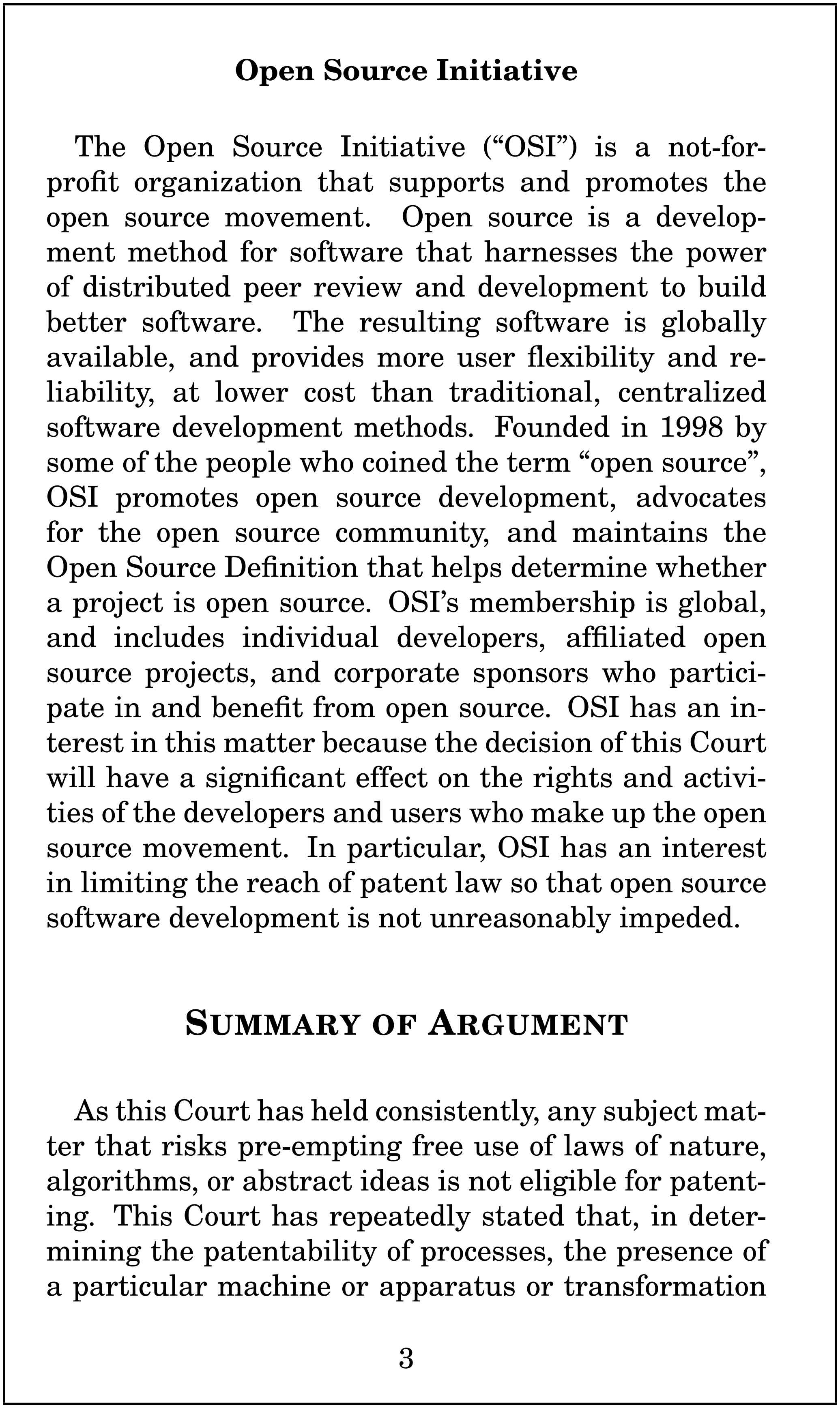
[1] Pursuant to Sup. Ct. R. 37.6, amici note that no counsel for a party authored this brief in whole or in part, and no counsel or party made a monetary contribution intended to fund the preparation or submission of this brief. No persons other than amici curiae and their counsel made a monetary contribution to its preparation or submission. Petitioners and Respondents have consented to the filing of this brief through blanket consent letters filed with the Clerk's Office.



free software developers and users SFLC represents. More specifically, SFLC has an interest in ensuring that limits are maintained on the reach of patent law so that free software development is not unreasonably and unnecessarily impeded.

Free Software Foundation

This brief is filed on behalf of the Free Software Foundation, a charitable corporation with its main offices in Boston, Massachusetts. The Foundation believes that people should be free to study, share and improve all the software they use and that this right is an essential freedom for users of computing. The Foundation has been working to achieve this goal since 1985 by directly developing and distributing, and by helping others to develop and distribute, software that is licensed on terms that permit all users to copy, modify and redistribute the works, so long as they give others the same freedoms to use, modify and redistribute in turn. The Foundation is the largest single contributor to the GNU operating system (used widely today in its GNU/Linux variant for computers from PCs to supercomputer clusters). The Foundation's GNU General Public License is the most widely used free software license, covering major components of the GNU operating system and hundreds of thousands of other computer programs used on hundreds of millions of computers around the world. The Foundation strongly rejects the use of patent law to control the making and distribution of software. It believes that the misapplication of patent law to computer software prevents the development, distribution and use of free/libre software, and therefore endangers users' control of their digital activities.

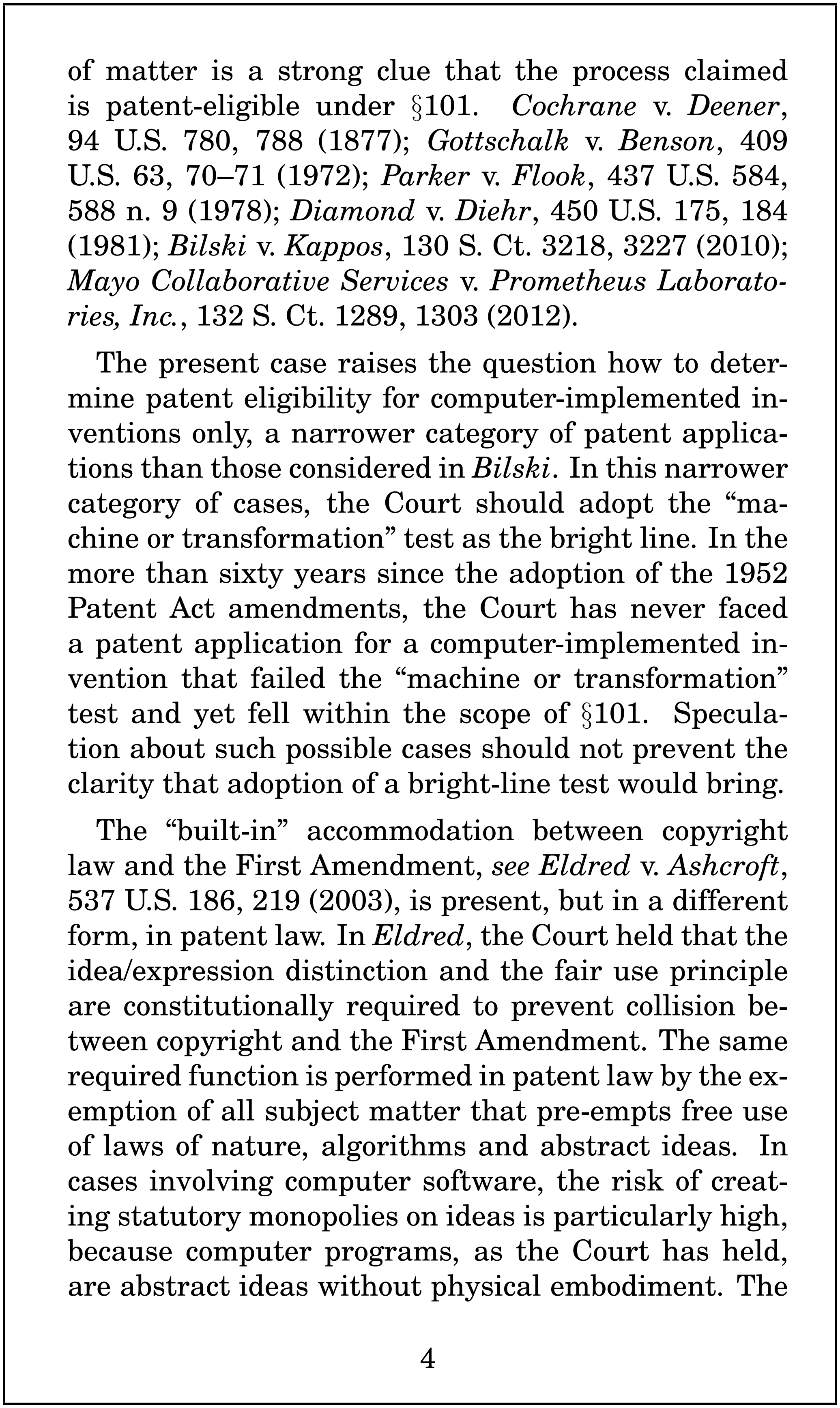


Open Source Initiative

The Open Source Initiative ("OSI") is a not-for-profit organization that supports and promotes the open source movement. Open source is a development method for software that harnesses the power of distributed peer review and development to build better software. The resulting software is globally available, and provides more user flexibility and reliability, at lower cost than traditional, centralized software development methods. Founded in 1998 by some of the people who coined the term "open source", OSI promotes open source development, advocates for the open source community, and maintains the Open Source Definition that helps determine whether a project is open source. OSI's membership is global, and includes individual developers, affiliated open source projects, and corporate sponsors who participate in and benefit from open source. OSI has an interest in this matter because the decision of this Court will have a significant effect on the rights and activities of the developers and users who make up the open source movement. In particular, OSI has an interest in limiting the reach of patent law so that open source software development is not unreasonably impeded.

Summary of Argument

As this Court has held consistently, any subject matter that risks pre-empting free use of laws of nature, algorithms, or abstract ideas is not eligible for patenting. This Court has repeatedly stated that, in determining the patentability of processes, the presence of a particular machine or apparatus or transformation



of matter is a strong clue that the process claimed is patent-eligible under §101. *Cochrane* v. *Deener*, 94 U.S. 780, 788 (1877); *Gottschalk* v. *Benson*, 409 U.S. 63, 70–71 (1972); *Parker* v. *Flook*, 437 U.S. 584, 588 n. 9 (1978); *Diamond* v. *Diehr*, 450 U.S. 175, 184 (1981); *Bilski* v. *Kappos*, 130 S. Ct. 3218, 3227 (2010); *Mayo Collaborative Services* v. *Prometheus Laboratories, Inc.*, 132 S. Ct. 1289, 1303 (2012).

The present case raises the question how to determine patent eligibility for computer-implemented inventions only, a narrower category of patent applications than those considered in *Bilski*. In this narrower category of cases, the Court should adopt the "machine or transformation" test as the bright line. In the more than sixty years since the adoption of the 1952 Patent Act amendments, the Court has never faced a patent application for a computer-implemented invention that failed the "machine or transformation" test and yet fell within the scope of §101. Speculation about such possible cases should not prevent the clarity that adoption of a bright-line test would bring.

The "built-in" accommodation between copyright law and the First Amendment, *see Eldred* v. *Ashcroft*, 537 U.S. 186, 219 (2003), is present, but in a different form, in patent law. In *Eldred*, the Court held that the idea/expression distinction and the fair use principle are constitutionally required to prevent collision between copyright and the First Amendment. The same required function is performed in patent law by the exemption of all subject matter that pre-empts free use of laws of nature, algorithms and abstract ideas. In cases involving computer software, the risk of creating statutory monopolies on ideas is particularly high, because computer programs, as the Court has held, are abstract ideas without physical embodiment. The

Court must construe the Patent Act to avoid constitutional infirmity, and it does so in this context by applying the "machine or transformation" test to such applications seeking statutory monopolies for computer-implemented inventions.

ARGUMENT

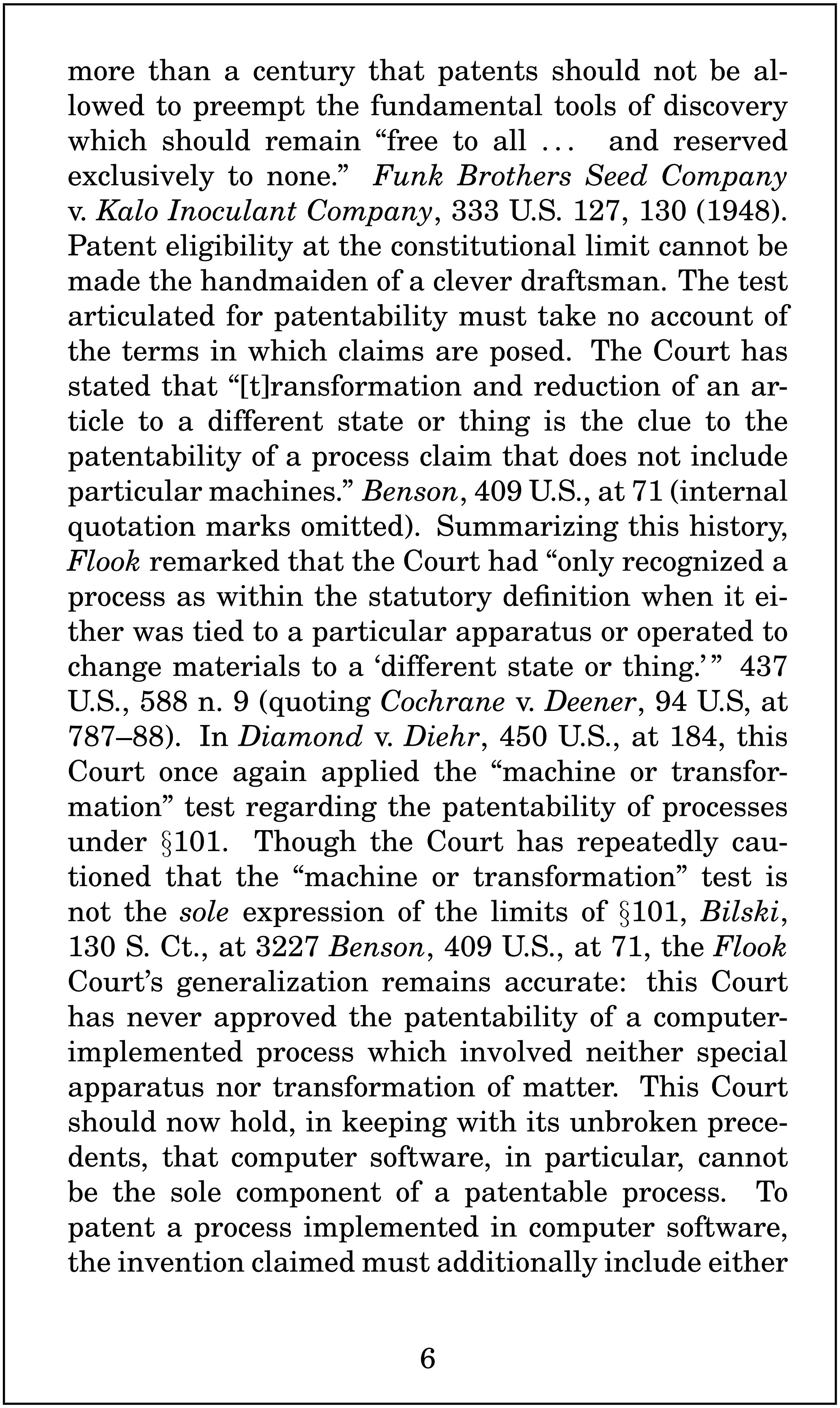
I. Processes Are Not Patentable If They Are Implemented Solely Through Computer Software, Without a Specialized Machine, or Transformation of Matter, As This Court Has Uniformly Held

As Justice Breyer noted in *Bilski*, 130 S. Ct., at 3258:

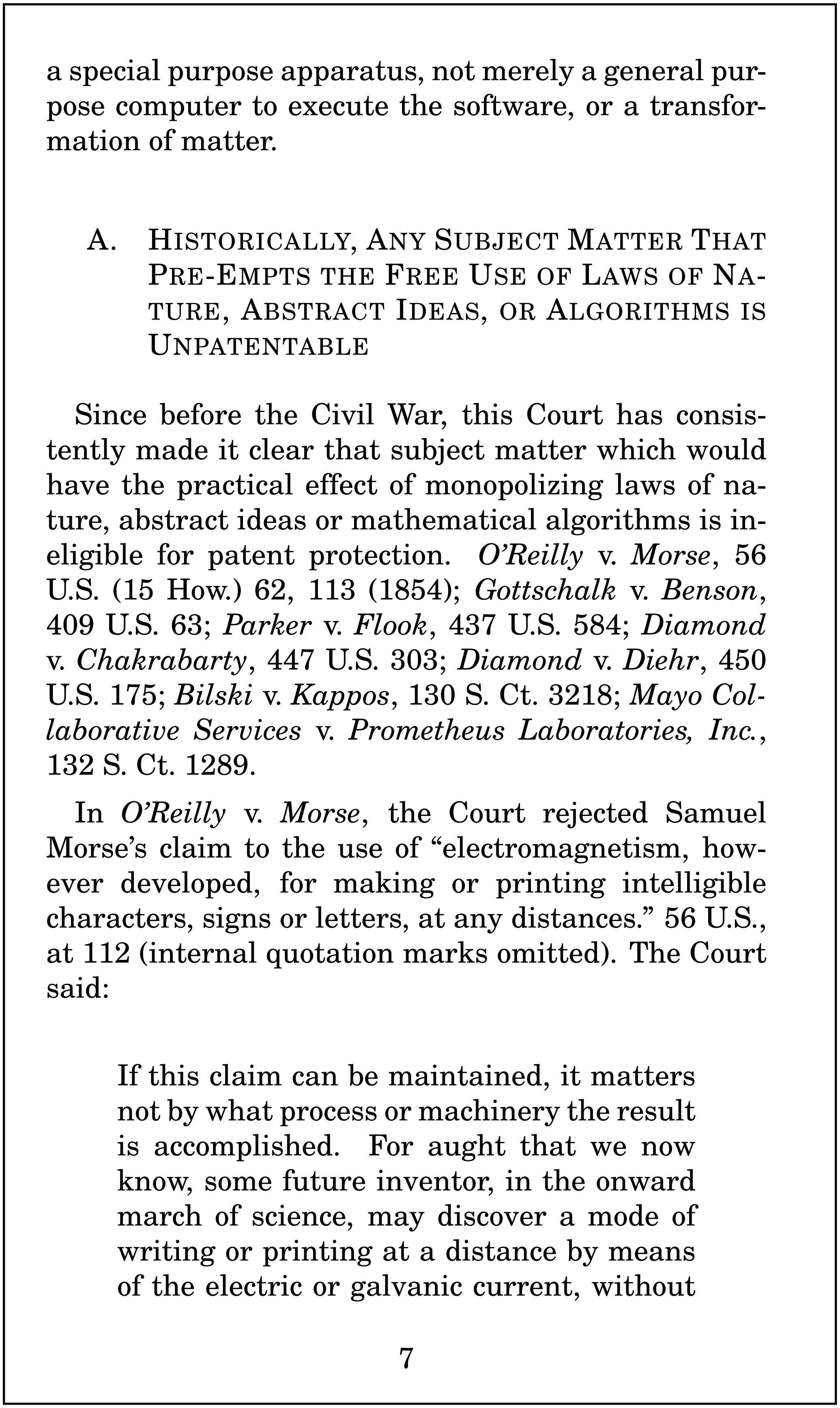
> [A]lthough the text of §101 is broad, it is not without limit. ... In particular, the Court has long held that '[p]henomena of nature, though just discovered, mental processes and abstract intellectual concepts are not patentable' under §101, since allowing individuals to patent these fundamental principles would 'wholly pre-empt' the public's access to the basic tools of scientific and technological work.

(internal citations omitted) (quoting *Benson*, 409 U.S., at 67, 72) (citing *Diehr*, 450 U.S., at 185; *Diamond* v. *Chakrabarty*, 447 U.S. 303, 309 (1980)).

In keeping with this principle, the Court has recognized in an unbroken series of cases extending over



more than a century that patents should not be allowed to preempt the fundamental tools of discovery which should remain "free to all ... and reserved exclusively to none." *Funk Brothers Seed Company* v. *Kalo Inoculant Company*, 333 U.S. 127, 130 (1948). Patent eligibility at the constitutional limit cannot be made the handmaiden of a clever draftsman. The test articulated for patentability must take no account of the terms in which claims are posed. The Court has stated that "[t]ransformation and reduction of an article to a different state or thing is the clue to the patentability of a process claim that does not include particular machines." *Benson*, 409 U.S., at 71 (internal quotation marks omitted). Summarizing this history, *Flook* remarked that the Court had "only recognized a process as within the statutory definition when it either was tied to a particular apparatus or operated to change materials to a 'different state or thing.'" 437 U.S., 588 n. 9 (quoting *Cochrane* v. *Deener*, 94 U.S, at 787–88). In *Diamond* v. *Diehr*, 450 U.S., at 184, this Court once again applied the "machine or transformation" test regarding the patentability of processes under §101. Though the Court has repeatedly cautioned that the "machine or transformation" test is not the *sole* expression of the limits of §101, *Bilski*, 130 S. Ct., at 3227 *Benson*, 409 U.S., at 71, the *Flook* Court's generalization remains accurate: this Court has never approved the patentability of a computer-implemented process which involved neither special apparatus nor transformation of matter. This Court should now hold, in keeping with its unbroken precedents, that computer software, in particular, cannot be the sole component of a patentable process. To patent a process implemented in computer software, the invention claimed must additionally include either



a special purpose apparatus, not merely a general purpose computer to execute the software, or a transformation of matter.

A. Historically, Any Subject Matter That Pre-Empts the Free Use of Laws of Nature, Abstract Ideas, or Algorithms is Unpatentable

Since before the Civil War, this Court has consistently made it clear that subject matter which would have the practical effect of monopolizing laws of nature, abstract ideas or mathematical algorithms is ineligible for patent protection. *O'Reilly* v. *Morse*, 56 U.S. (15 How.) 62, 113 (1854); *Gottschalk* v. *Benson*, 409 U.S. 63; *Parker* v. *Flook*, 437 U.S. 584; *Diamond* v. *Chakrabarty*, 447 U.S. 303; *Diamond* v. *Diehr*, 450 U.S. 175; *Bilski* v. *Kappos*, 130 S. Ct. 3218; *Mayo Collaborative Services* v. *Prometheus Laboratories, Inc.*, 132 S. Ct. 1289.

In *O'Reilly* v. *Morse*, the Court rejected Samuel Morse's claim to the use of "electromagnetism, however developed, for making or printing intelligible characters, signs or letters, at any distances." 56 U.S., at 112 (internal quotation marks omitted). The Court said:

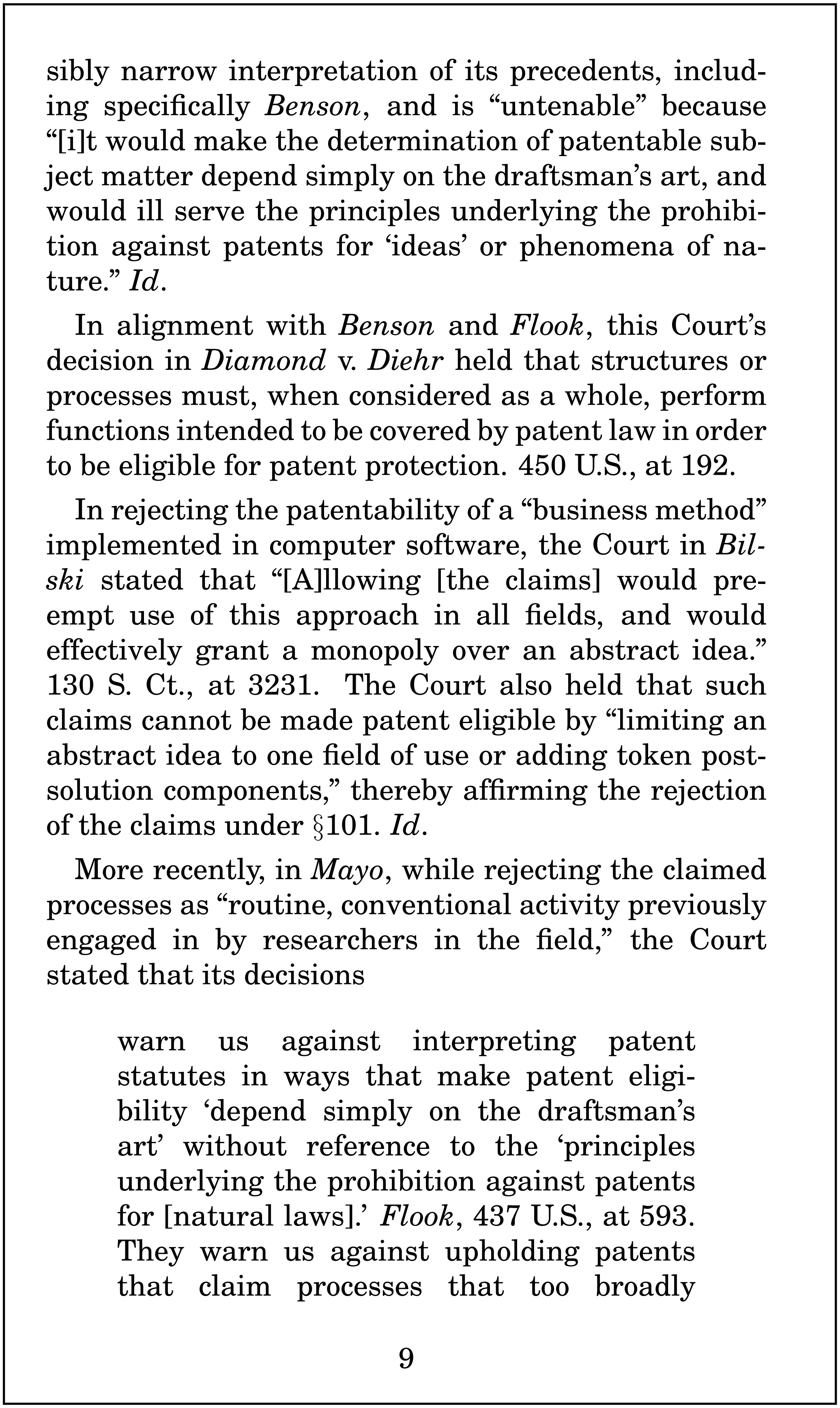
> If this claim can be maintained, it matters not by what process or machinery the result is accomplished. For aught that we now know, some future inventor, in the onward march of science, may discover a mode of writing or printing at a distance by means of the electric or galvanic current, without

> using any part of the process or combination set forth in the plaintiff's specification. His invention may be less complicated — less liable to get out of order — less expensive in construction, and in its operation. But yet, if it is covered by this patent, the inventor could not use it, nor the public have the benefit of it, without the permission of this patentee.

Id. at 113.

In *Benson*, this Court considered the claim to a method for converting numerical information from binary-coded decimal numbers into pure binary numbers, for use in programming conventional general-purpose digital computers. The Court concluded that "[t]he mathematical formula involved here has no substantial practical application except in connection with a digital computer, which means that if the judgment below is affirmed, the patent would wholly pre-empt the mathematical formula involved and in practical effect would be a patent on the algorithm itself." 409 U.S., at 71–72. Accordingly the claims were held ineligible under §101.

In *Parker* v. *Flook*, the Court held that to be eligible for patent protection, "[t]he process itself, not merely the mathematical algorithm, must be new and useful." 437 U.S., at 591; *see also*, *Funk Brothers*, 333 U.S., at 130. The Court further stated in *Flook* that it is "incorrect[] [to] assume[] that if a process application implements a principle in some specific fashion, it automatically falls within the patentable subject matter of §101." 437 U.S., at 593. This Court explained that such an assumption is based on an impermis-



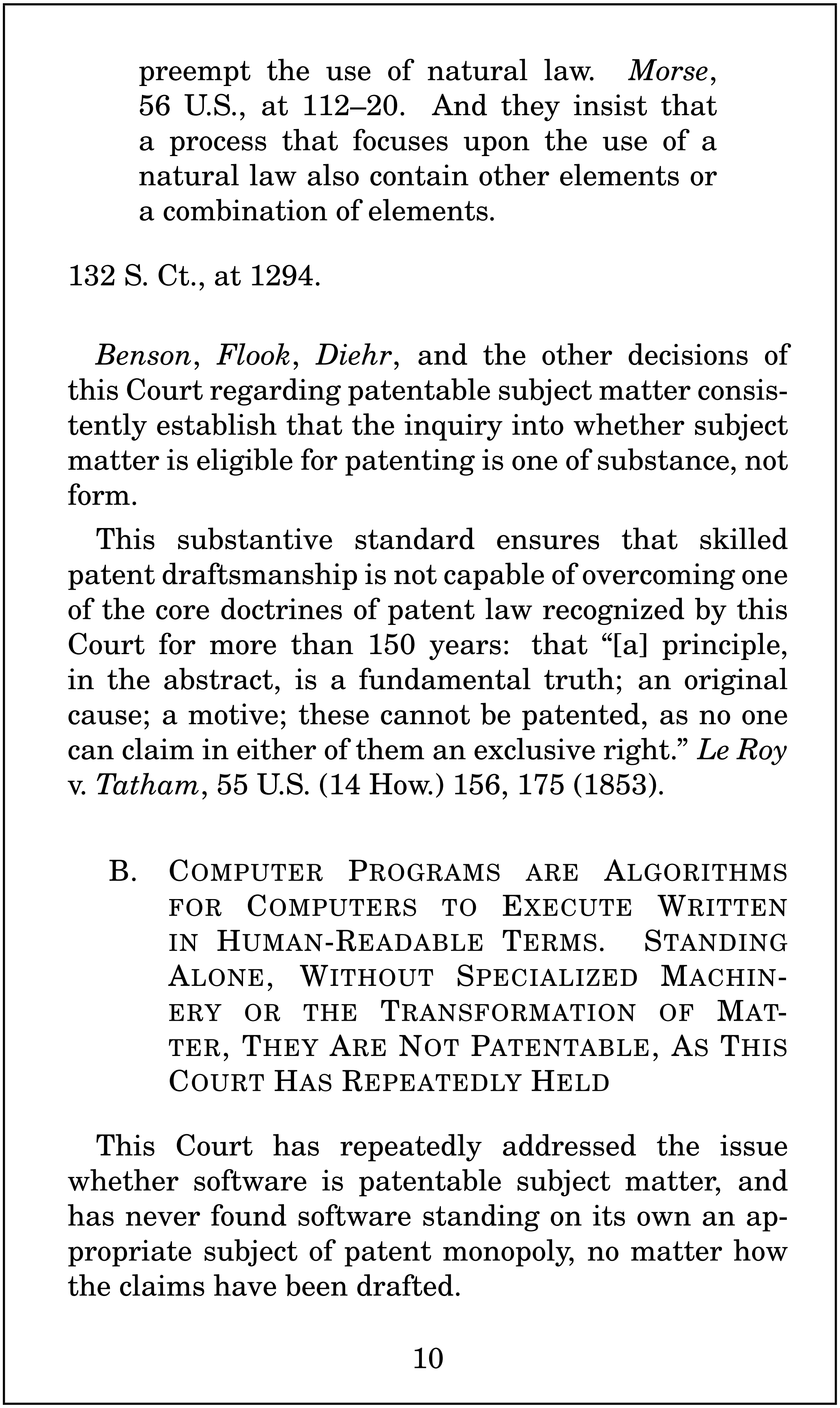
sibly narrow interpretation of its precedents, including specifically *Benson*, and is "untenable" because "[i]t would make the determination of patentable subject matter depend simply on the draftsman's art, and would ill serve the principles underlying the prohibition against patents for 'ideas' or phenomena of nature." *Id*.

In alignment with *Benson* and *Flook*, this Court's decision in *Diamond* v. *Diehr* held that structures or processes must, when considered as a whole, perform functions intended to be covered by patent law in order to be eligible for patent protection. 450 U.S., at 192.

In rejecting the patentability of a "business method" implemented in computer software, the Court in *Bilski* stated that "[A]llowing [the claims] would preempt use of this approach in all fields, and would effectively grant a monopoly over an abstract idea." 130 S. Ct., at 3231. The Court also held that such claims cannot be made patent eligible by "limiting an abstract idea to one field of use or adding token post-solution components," thereby affirming the rejection of the claims under §101. *Id*.

More recently, in *Mayo*, while rejecting the claimed processes as "routine, conventional activity previously engaged in by researchers in the field," the Court stated that its decisions

> warn us against interpreting patent statutes in ways that make patent eligibility 'depend simply on the draftsman's art' without reference to the 'principles underlying the prohibition against patents for [natural laws].' *Flook*, 437 U.S., at 593. They warn us against upholding patents that claim processes that too broadly



> preempt the use of natural law. *Morse*, 56 U.S., at 112–20. And they insist that a process that focuses upon the use of a natural law also contain other elements or a combination of elements.

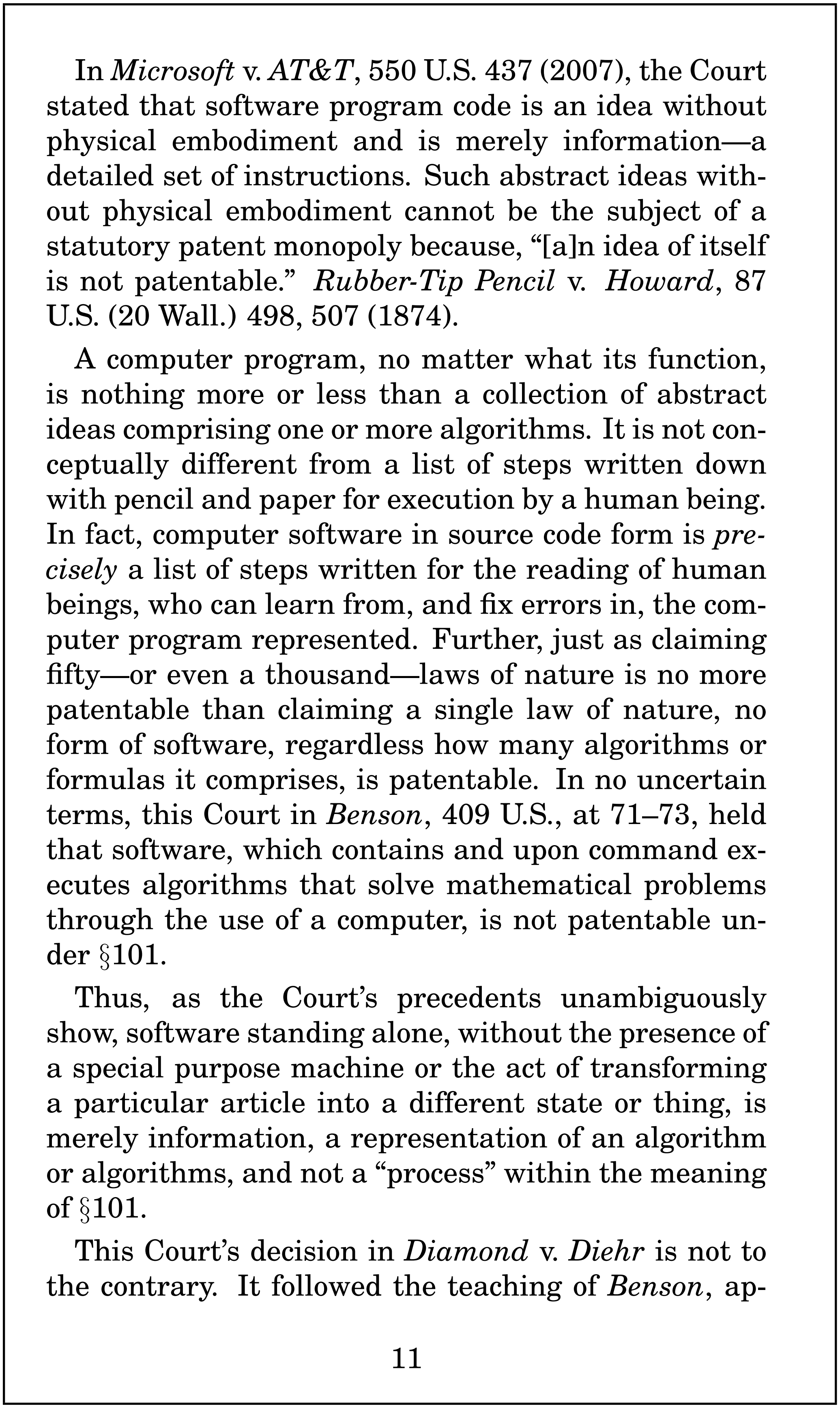
132 S. Ct., at 1294.

Benson, *Flook*, *Diehr*, and the other decisions of this Court regarding patentable subject matter consistently establish that the inquiry into whether subject matter is eligible for patenting is one of substance, not form.

This substantive standard ensures that skilled patent draftsmanship is not capable of overcoming one of the core doctrines of patent law recognized by this Court for more than 150 years: that "[a] principle, in the abstract, is a fundamental truth; an original cause; a motive; these cannot be patented, as no one can claim in either of them an exclusive right." *Le Roy* v. *Tatham*, 55 U.S. (14 How.) 156, 175 (1853).

B. Computer Programs are Algorithms for Computers to Execute Written in Human-Readable Terms. Standing Alone, Without Specialized Machinery or the Transformation of Matter, They Are Not Patentable, As This Court Has Repeatedly Held

This Court has repeatedly addressed the issue whether software is patentable subject matter, and has never found software standing on its own an appropriate subject of patent monopoly, no matter how the claims have been drafted.



In *Microsoft* v. *AT&T*, 550 U.S. 437 (2007), the Court stated that software program code is an idea without physical embodiment and is merely information—a detailed set of instructions. Such abstract ideas without physical embodiment cannot be the subject of a statutory patent monopoly because, "[a]n idea of itself is not patentable." *Rubber-Tip Pencil* v. *Howard*, 87 U.S. (20 Wall.) 498, 507 (1874).

A computer program, no matter what its function, is nothing more or less than a collection of abstract ideas comprising one or more algorithms. It is not conceptually different from a list of steps written down with pencil and paper for execution by a human being. In fact, computer software in source code form is *precisely* a list of steps written for the reading of human beings, who can learn from, and fix errors in, the computer program represented. Further, just as claiming fifty—or even a thousand—laws of nature is no more patentable than claiming a single law of nature, no form of software, regardless how many algorithms or formulas it comprises, is patentable. In no uncertain terms, this Court in *Benson*, 409 U.S., at 71–73, held that software, which contains and upon command executes algorithms that solve mathematical problems through the use of a computer, is not patentable under §101.

Thus, as the Court's precedents unambiguously show, software standing alone, without the presence of a special purpose machine or the act of transforming a particular article into a different state or thing, is merely information, a representation of an algorithm or algorithms, and not a "process" within the meaning of §101.

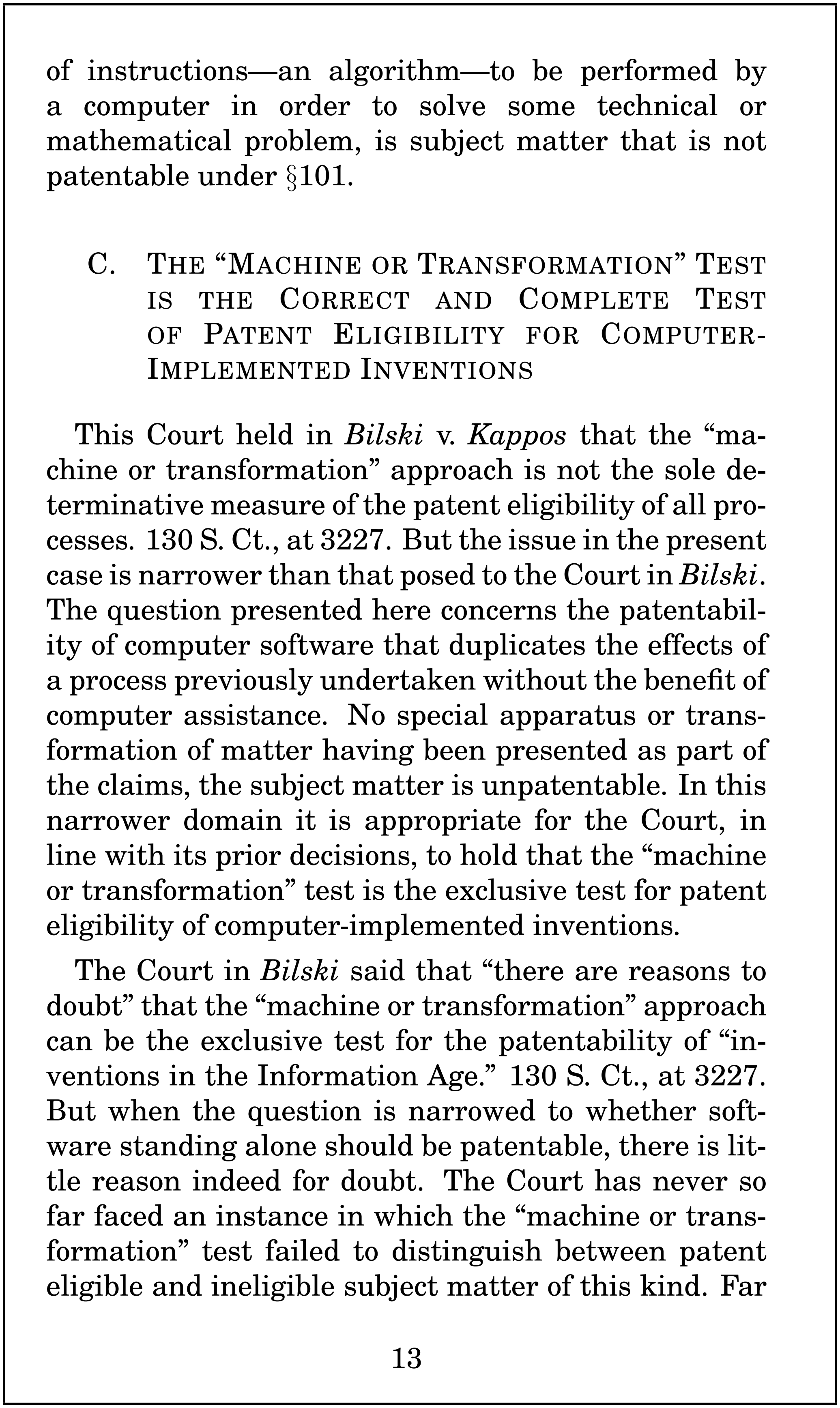
This Court's decision in *Diamond* v. *Diehr* is not to the contrary. It followed the teaching of *Benson*, applied in substance the "machine or transformation" test, and determined that the invention before the Court was not substantially the software, but rather the totality of an "industrial process for the molding of rubber products," which was undeniably included within the realm of patentable subject matter. 450 U.S., at 191–93. Had the applicant sought to claim the software used in that process by itself, however, the Court would surely have found it to be unpatentable subject matter just as it had in *Benson*. As the *Diehr* Court observed:

> [W]hen a claim recites a mathematical formula (or scientific principle or phenomenon of nature), an inquiry must be made into whether the claim is seeking patent protection for that formula in the abstract. A mathematical formula as such is not accorded the protection of patent laws, and this principle cannot be circumvented by attempting to limit the use of the formula to a particular technological environment.

450 U.S., at 191 (internal citations omitted).

This result—which makes software describing a portion of the solution to a practical problem unpatentable on its own, outside the real-world context of the problem and its solution—is not only in accord with the rest of this Court's patent jurisprudence, it is also the best way to protect innovation in software, and the only way that fully comports with both Article I, §8 and the First Amendment.

Thus, this Court's precedent repeatedly sets out that software, which is nothing more than a set



of instructions—an algorithm—to be performed by a computer in order to solve some technical or mathematical problem, is subject matter that is not patentable under §101.

C. The "Machine or Transformation" Test is the Correct and Complete Test of Patent Eligibility for Computer-Implemented Inventions

This Court held in *Bilski* v. *Kappos* that the "machine or transformation" approach is not the sole determinative measure of the patent eligibility of all processes. 130 S. Ct., at 3227. But the issue in the present case is narrower than that posed to the Court in *Bilski*. The question presented here concerns the patentability of computer software that duplicates the effects of a process previously undertaken without the benefit of computer assistance. No special apparatus or transformation of matter having been presented as part of the claims, the subject matter is unpatentable. In this narrower domain it is appropriate for the Court, in line with its prior decisions, to hold that the "machine or transformation" test is the exclusive test for patent eligibility of computer-implemented inventions.

The Court in *Bilski* said that "there are reasons to doubt" that the "machine or transformation" approach can be the exclusive test for the patentability of "inventions in the Information Age." 130 S. Ct., at 3227. But when the question is narrowed to whether software standing alone should be patentable, there is little reason indeed for doubt. The Court has never so far faced an instance in which the "machine or transformation" test failed to distinguish between patent eligible and ineligible subject matter of this kind. Far from being a source of uncertainty, as the *Bilski* Court suggested it might be, *Id.*, the “machine or transformation” test would provide substantial certainty now lacking, by reinforcing the teaching of the unbroken precedent of 150 years.

No doubt the pace of change in the area of information technology is rapid. As we show below, the real lesson of contemporary technological development, however, is that patenting has had no positive effect on innovation in software. Uncertainty about what can be patented, on the other hand, has given rise to enormously wasteful litigation. But whatever the pace of innovation, it is unlikely to disclose what has not yet appeared since the beginning of the Information Age: a case in which software standing alone, that fails the “machine or transformation” test, nonetheless is patentable subject matter. In the unlikely event that such a *rara avis* is observed in future, the Court can modify or add to the test. In the meantime, the advantages of certainty that would accrue from the adoption of a clear, bright-line test that cannot be defeated by mere cleverness of draftsmanship would far outweigh the speculative concerns expressed by the Court in *Bilski*.

II. Adhering to the Court’s Previous Decisions on Patentability of Software Standing Alone Does Not Imperil the Pace of Software Innovation

The nature of software, like mathematics or basic scientific research, is that innovation is best produced by free sharing. History shows that innovation in software over the last generation has occurred first in communities of free sharing, where patenting has been systematically discouraged.

A. Innovation in Software, Like Innovation in Mathematics, is Encouraged by Scientific Processes of Free Sharing and Open Publication, Not by Granting State-Issued Monopolies on Ideas

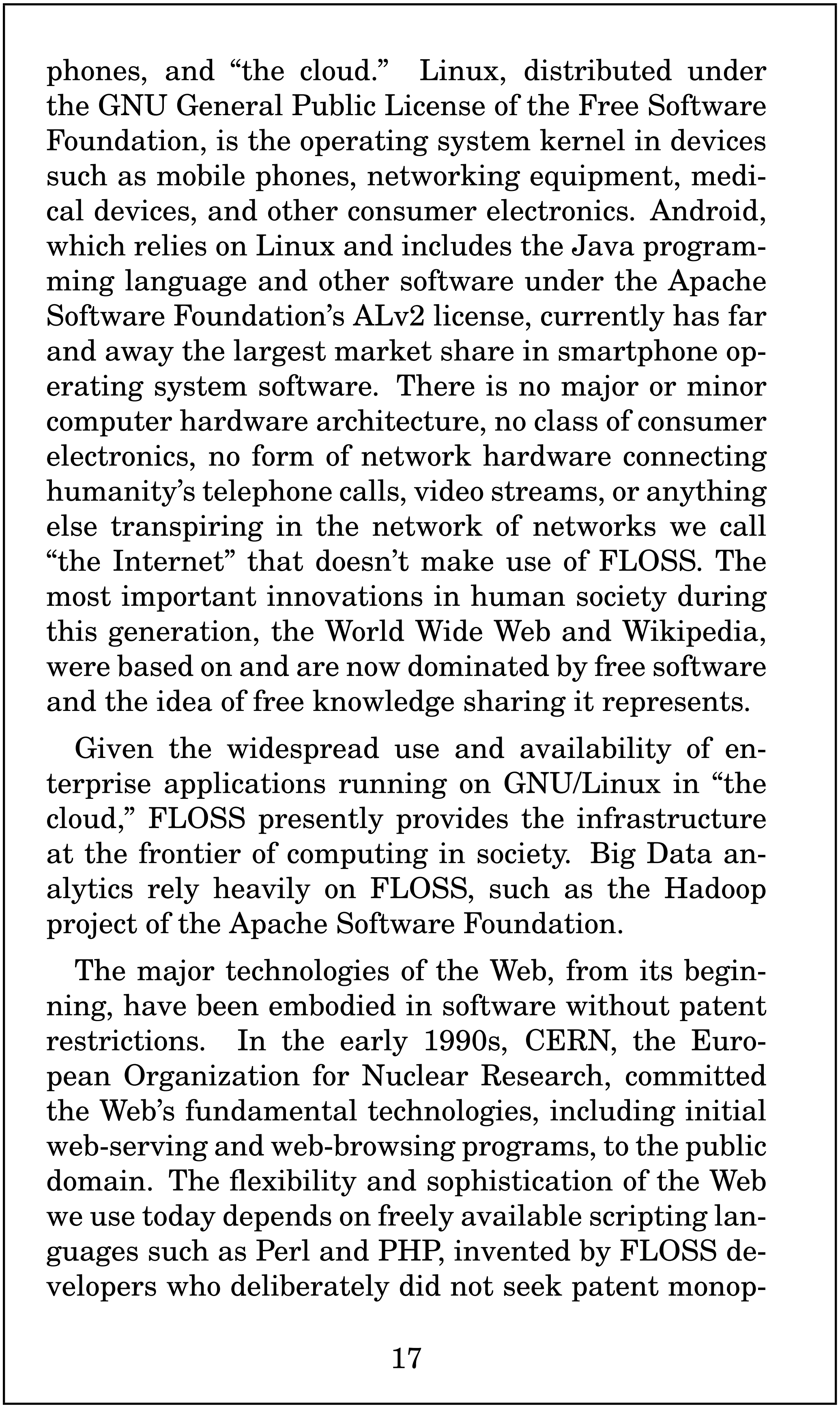
If mathematics were patentable, there would be less mathematical innovation. Only those who were rich enough to pay royalties, or who benefited from subsidization by government, or who were willing to sign over the value of their ideas to someone richer and more powerful than themselves, would be permitted access to the world of abstract mathematical ideas. Theorems build upon theorems, and so the contributions of those who could not pay rent—and all the further improvements based upon those contributions—would be lost.

The principle that innovation is made possible by the free exchange of ideas is not recent, and is not limited to software. Indeed, our constitutional system of free expression since Thomas Jefferson is based on the recognition that control of ideas by power has never produced more ideas than their free and unrestricted circulation. The history of western science since the 17th century is one long testament to this truth, and it is that very history which gave rise to the patent system, whose exclusion of "abstract ideas," "laws of nature," and "algorithms" is as much a recognition of the principle as is the basic constitutional policy of offering temporary legal benefits in return for prompt and complete disclosure of technological discoveries to the public.

B. The History of the Free Software Movement and the Worldwide Adoption of Free and Open Source Software by Industry Shows That Patenting Software Has Not Contributed to the Important Software Innovations of the Last Generation

For more than a quarter century, beginning with a few stalwart thinkers and exponentially increasing in size and influence, a movement to build computer software by sharing—treating software programming languages like mathematical notation, for the expression of abstract ideas to be studied, improved, and shared again—has revolutionized the production of software around the world. The "free software movement," and the developers of "open source software" (collectively described hereinafter as "FLOSS developers") believe, like this Court, that computer software expresses abstract ideas. FLOSS developers therefore conclude that the ideas themselves will grow best if left most free to be learned and improved by all. Their conviction has been shared by hundreds of thousands—soon millions—of programmers around the world, who have devoted their skills to making new and innovative software through the social process that for centuries has been the heart of Western science: "share and share alike."

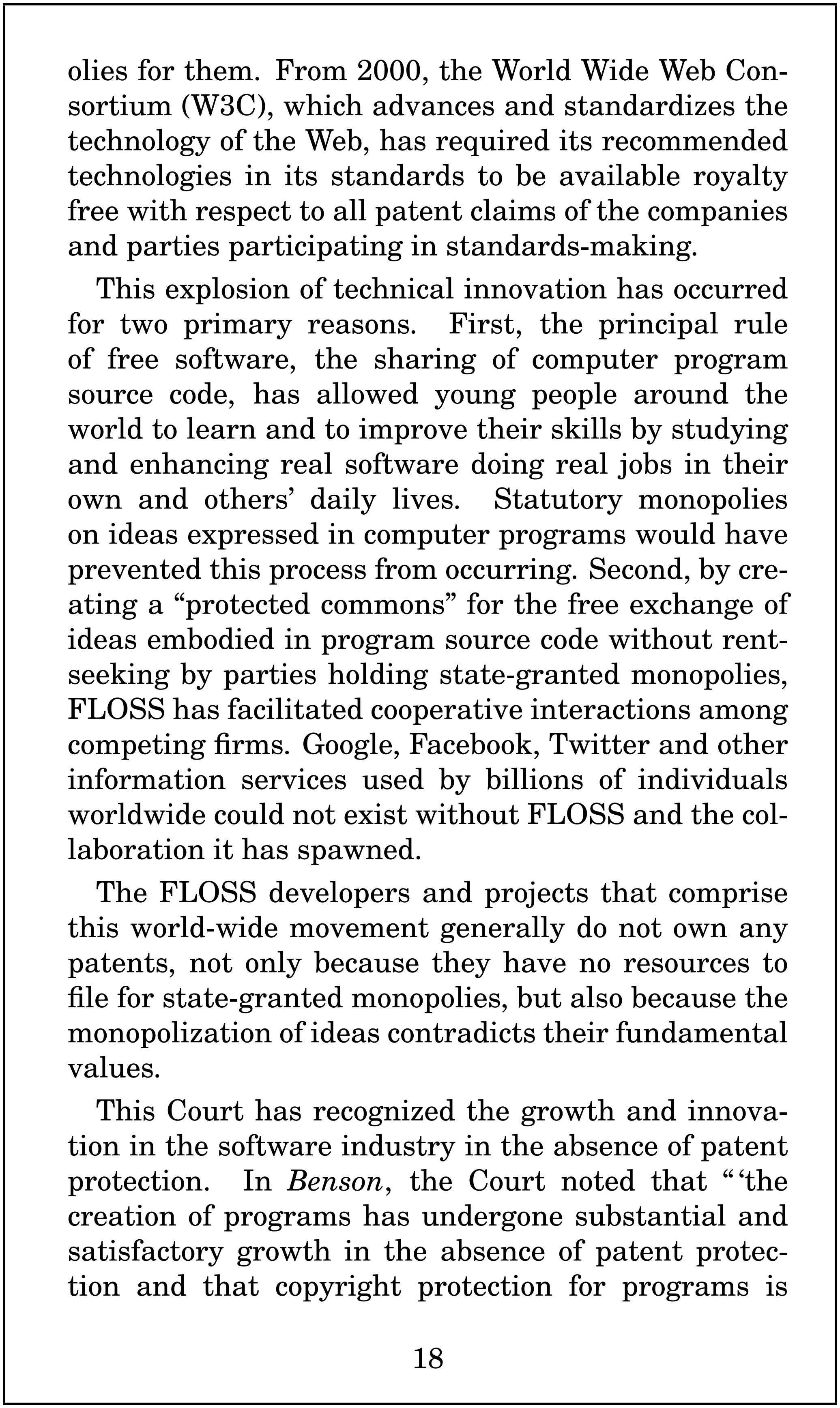
FLOSS has become the single most influential body of software around the world. In the more than twenty years of its existence, FLOSS has taken the world by storm and has driven the majority of the world's technological advancement in computer programming. FLOSS lives under the hood of it all—from desktops and servers, to laptops, netbooks, smart-



phones, and “the cloud.” Linux, distributed under the GNU General Public License of the Free Software Foundation, is the operating system kernel in devices such as mobile phones, networking equipment, medical devices, and other consumer electronics. Android, which relies on Linux and includes the Java programming language and other software under the Apache Software Foundation’s ALv2 license, currently has far and away the largest market share in smartphone operating system software. There is no major or minor computer hardware architecture, no class of consumer electronics, no form of network hardware connecting humanity’s telephone calls, video streams, or anything else transpiring in the network of networks we call “the Internet” that doesn’t make use of FLOSS. The most important innovations in human society during this generation, the World Wide Web and Wikipedia, were based on and are now dominated by free software and the idea of free knowledge sharing it represents.

Given the widespread use and availability of enterprise applications running on GNU/Linux in “the cloud,” FLOSS presently provides the infrastructure at the frontier of computing in society. Big Data analytics rely heavily on FLOSS, such as the Hadoop project of the Apache Software Foundation.

The major technologies of the Web, from its beginning, have been embodied in software without patent restrictions. In the early 1990s, CERN, the European Organization for Nuclear Research, committed the Web’s fundamental technologies, including initial web-serving and web-browsing programs, to the public domain. The flexibility and sophistication of the Web we use today depends on freely available scripting languages such as Perl and PHP, invented by FLOSS developers who deliberately did not seek patent monop-

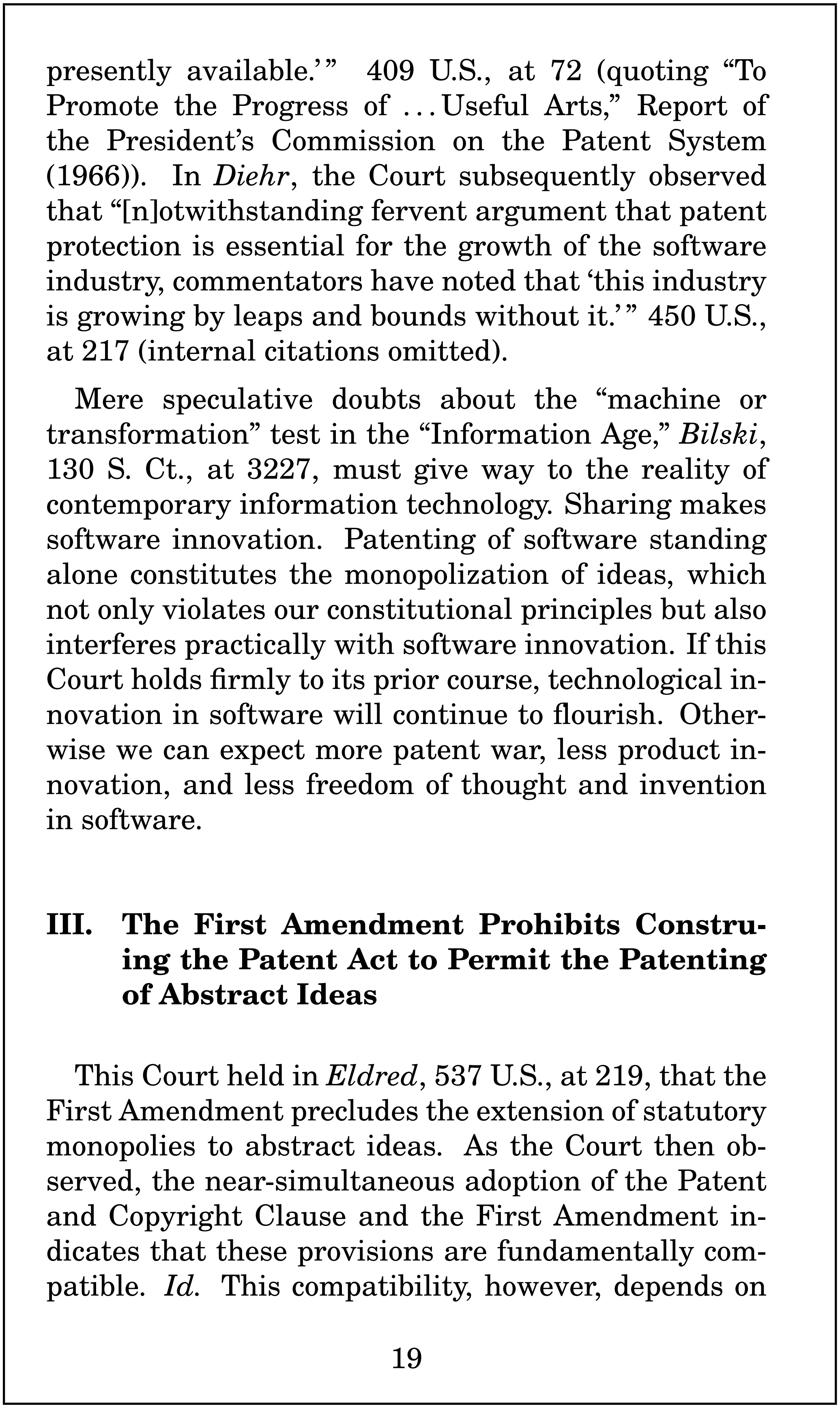


olies for them. From 2000, the World Wide Web Consortium (W3C), which advances and standardizes the technology of the Web, has required its recommended technologies in its standards to be available royalty free with respect to all patent claims of the companies and parties participating in standards-making.

This explosion of technical innovation has occurred for two primary reasons. First, the principal rule of free software, the sharing of computer program source code, has allowed young people around the world to learn and to improve their skills by studying and enhancing real software doing real jobs in their own and others' daily lives. Statutory monopolies on ideas expressed in computer programs would have prevented this process from occurring. Second, by creating a "protected commons" for the free exchange of ideas embodied in program source code without rent-seeking by parties holding state-granted monopolies, FLOSS has facilitated cooperative interactions among competing firms. Google, Facebook, Twitter and other information services used by billions of individuals worldwide could not exist without FLOSS and the collaboration it has spawned.

The FLOSS developers and projects that comprise this world-wide movement generally do not own any patents, not only because they have no resources to file for state-granted monopolies, but also because the monopolization of ideas contradicts their fundamental values.

This Court has recognized the growth and innovation in the software industry in the absence of patent protection. In *Benson*, the Court noted that "'the creation of programs has undergone substantial and satisfactory growth in the absence of patent protection and that copyright protection for programs is



presently available.'" 409 U.S., at 72 (quoting "To Promote the Progress of . . . Useful Arts," Report of the President's Commission on the Patent System (1966)). In *Diehr*, the Court subsequently observed that "[n]otwithstanding fervent argument that patent protection is essential for the growth of the software industry, commentators have noted that 'this industry is growing by leaps and bounds without it.'" 450 U.S., at 217 (internal citations omitted).

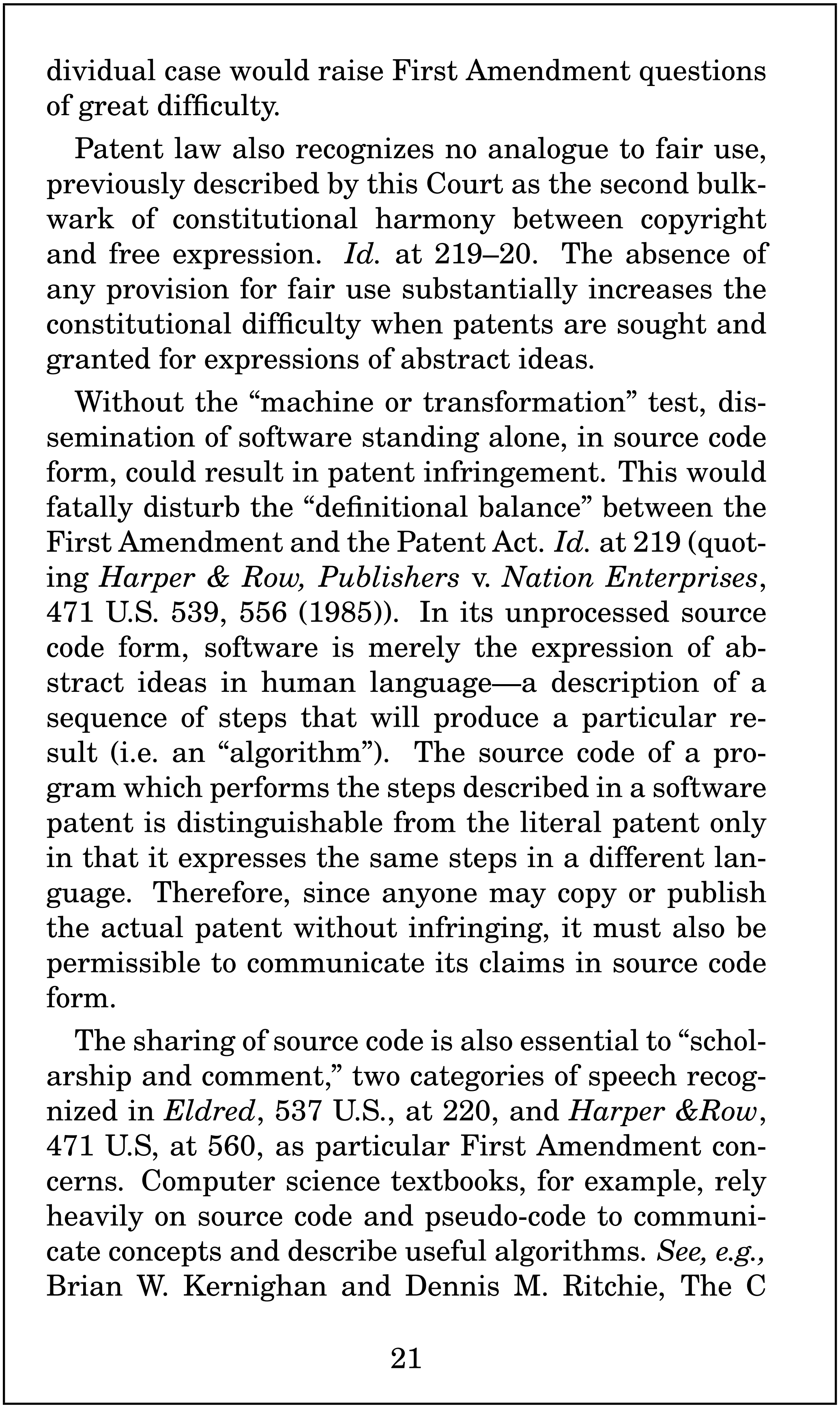
Mere speculative doubts about the "machine or transformation" test in the "Information Age," *Bilski*, 130 S. Ct., at 3227, must give way to the reality of contemporary information technology. Sharing makes software innovation. Patenting of software standing alone constitutes the monopolization of ideas, which not only violates our constitutional principles but also interferes practically with software innovation. If this Court holds firmly to its prior course, technological innovation in software will continue to flourish. Otherwise we can expect more patent war, less product innovation, and less freedom of thought and invention in software.

III. The First Amendment Prohibits Construing the Patent Act to Permit the Patenting of Abstract Ideas

This Court held in *Eldred*, 537 U.S., at 219, that the First Amendment precludes the extension of statutory monopolies to abstract ideas. As the Court then observed, the near-simultaneous adoption of the Patent and Copyright Clause and the First Amendment indicates that these provisions are fundamentally compatible. *Id.* This compatibility, however, depends on a construction of the patent and copyright acts that preserves First Amendment principles, including the freedom to communicate any "idea, theory, and fact." *Id.*

Eldred identified two mechanisms in copyright law that are necessary to accommodate this principle. First, the idea/expression dichotomy limits copyright's monopoly to an author's expression, leaving ideas "instantly available for public exploitation." *Id.* Second, the fair use doctrine allows the public to use even copyrighted expression for some purposes, "such as criticism, comment, news reporting, teaching. . . , scholarship, or research." *Id.* at 220.

Patent statutes, which depend on the same constitutional grant of authority as copyright statutes, are similarly limited by the First Amendment. *See id.* at 201 ("Because the Clause empowering Congress to confer copyrights also authorizes patents, congressional practice with respect to patents informs our inquiry"). The presence of an unwavering exemption for abstract ideas reconciles patent law with the First Amendment in a fashion similar to the idea/expression dichotomy's crucial role in reconciling copyright and freedom of speech. The presence of a limiting principle is even more necessary with respect to patent law than with respect to copyright, because, as the Court observed in *Eldred*, "the grant of a patent . . . prevent[s] full use by others of the inventor's knowledge." *Id.* at 217 (internal citation omitted). Patents can and do limit the application of knowledge to produce a new machine or to transform an article into a different state or thing, but they cannot constitutionally limit the communication of knowledge or ideas. *Eldred* teaches that, without this limitation, determining the scope of patent eligibility in each in-



dividual case would raise First Amendment questions of great difficulty.

Patent law also recognizes no analogue to fair use, previously described by this Court as the second bulkwark of constitutional harmony between copyright and free expression. *Id.* at 219–20. The absence of any provision for fair use substantially increases the constitutional difficulty when patents are sought and granted for expressions of abstract ideas.

Without the "machine or transformation" test, dissemination of software standing alone, in source code form, could result in patent infringement. This would fatally disturb the "definitional balance" between the First Amendment and the Patent Act. *Id.* at 219 (quoting *Harper & Row, Publishers* v. *Nation Enterprises*, 471 U.S. 539, 556 (1985)). In its unprocessed source code form, software is merely the expression of abstract ideas in human language—a description of a sequence of steps that will produce a particular result (i.e. an "algorithm"). The source code of a program which performs the steps described in a software patent is distinguishable from the literal patent only in that it expresses the same steps in a different language. Therefore, since anyone may copy or publish the actual patent without infringing, it must also be permissible to communicate its claims in source code form.

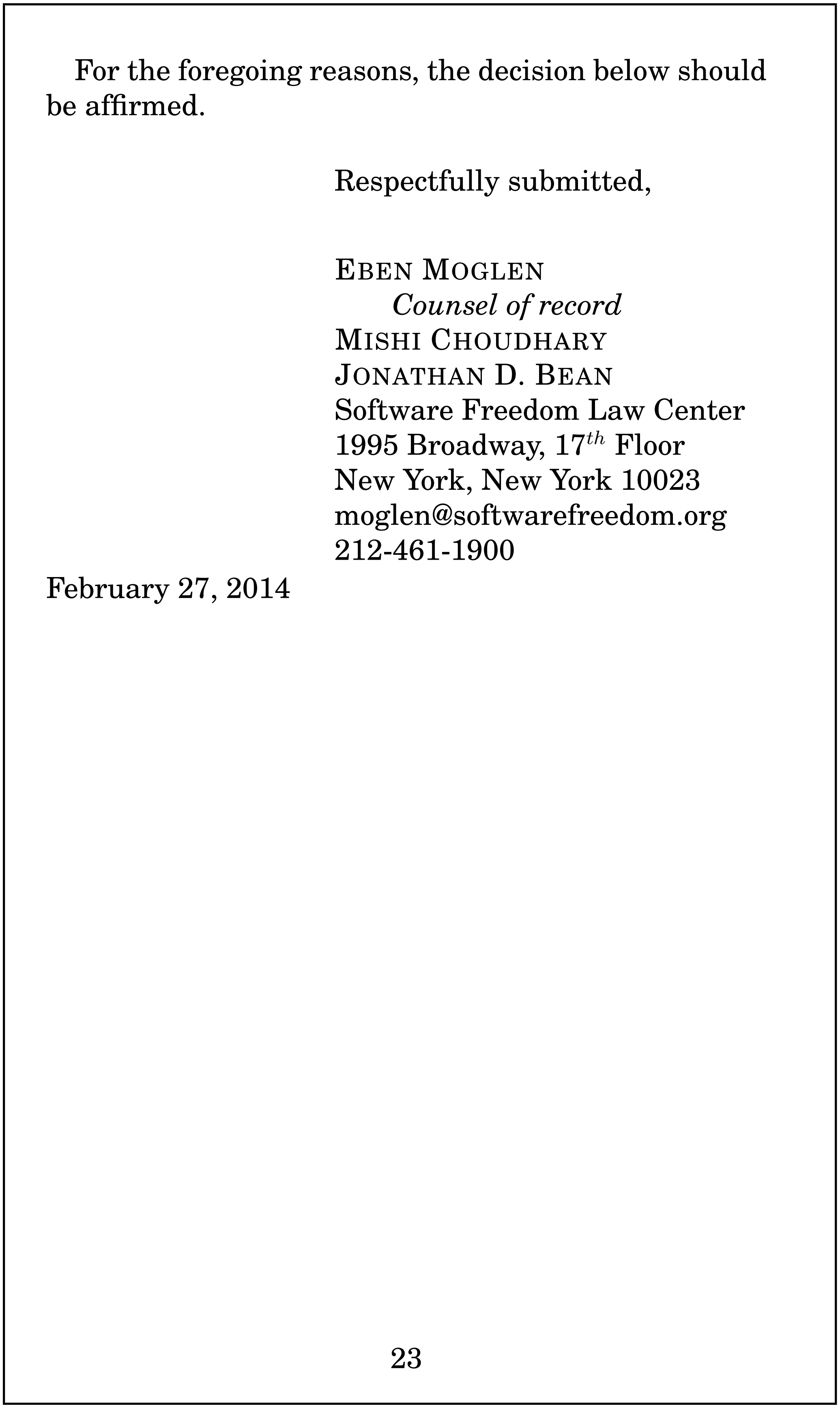
The sharing of source code is also essential to "scholarship and comment," two categories of speech recognized in *Eldred*, 537 U.S., at 220, and *Harper &Row*, 471 U.S, at 560, as particular First Amendment concerns. Computer science textbooks, for example, rely heavily on source code and pseudo-code to communicate concepts and describe useful algorithms. *See, e.g.,* Brian W. Kernighan and Dennis M. Ritchie, The C

Programming Language (Prentice Hall 1978). Likewise, computer science students are often required to express their answers to test questions in a real or hypothetical programming language. And without the use of source code, it is difficult for developers to comment on whether an idea can be implemented, to comment on an algorithm's performance, or to suggest improvements.

The "machine or transformation" test serves the purpose of securing accommodation with the First Amendment by ensuring that patent claims on computer-implemented inventions cannot be comprised solely of ideas communicated in computer program code. By requiring a physical special-purpose apparatus or a material transformation, the test implements a construction of §101 that automatically avoids conflict with the First Amendment. If the "machine or transformation" test is *not* the exclusive delimitation of §101 as applied to computer-implemented inventions, what alternative proposal do petitioners and their *amici* advance to avoid First Amendment problems?

CONCLUSION

The "machine or transformation" test evolved over 150 years of this Court's jurisprudence should be affirmed as the necessary criterion for the patenting of inventions implemented in software.



For the foregoing reasons, the decision below should be affirmed.

Respectfully submitted,

Eben Moglen
Counsel of record
Mishi Choudhary
Jonathan D. Bean
Software Freedom Law Center
1995 Broadway, 17th Floor
New York, New York 10023
moglen@softwarefreedom.org
212-461-1900

February 27, 2014

Part III

Automotive FOSS

Chapter 14

Automotive Software Governance and Copyleft (Shuttleworth, Moglen)

Automotive Software Governance and Copyleft

Mark Shuttleworth[*] Eben Moglen[†]

October 12, 2018

Introduction

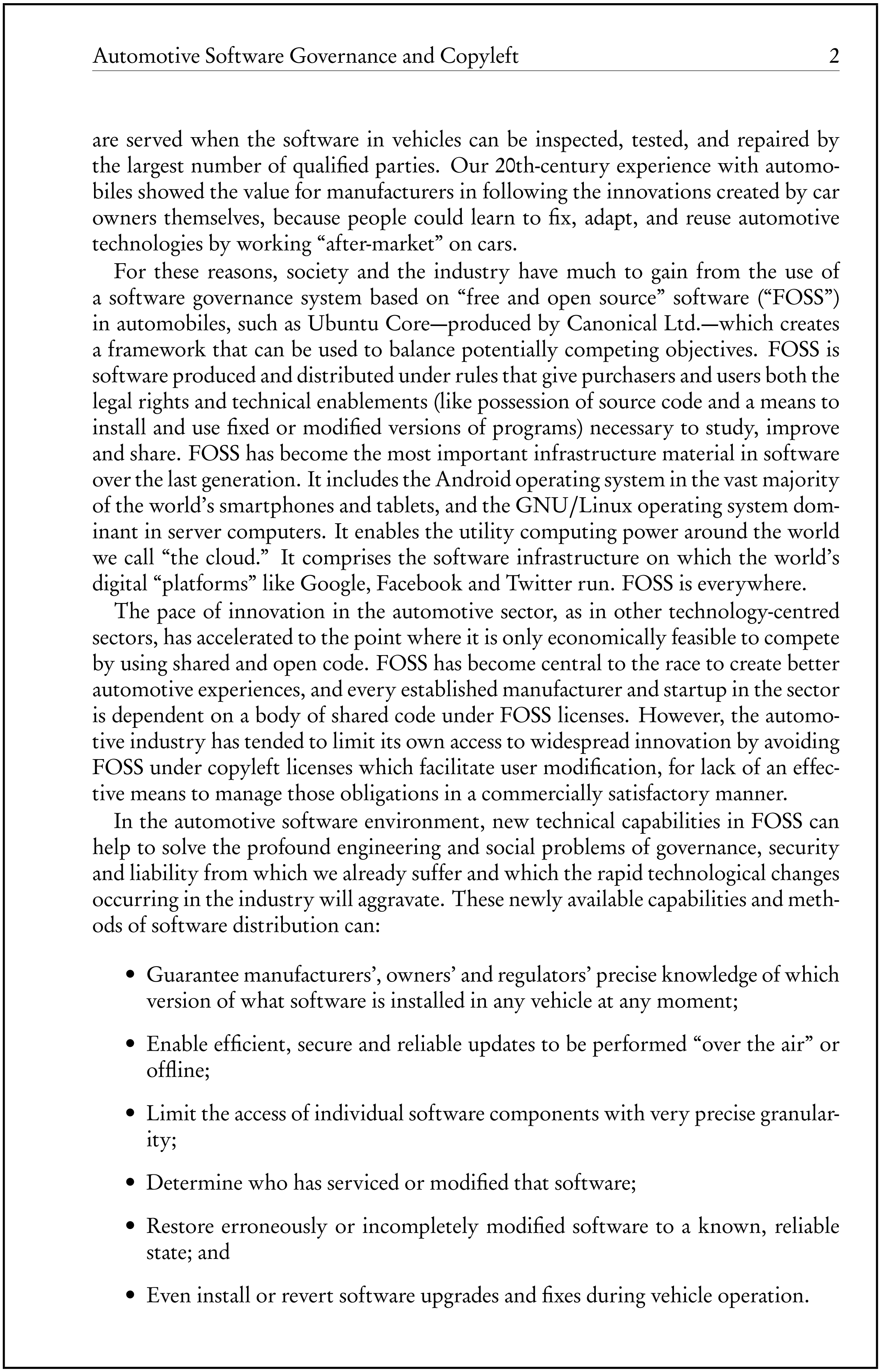
Our purpose in writing this paper is to show how new capabilities in the free and open source software stack enable highly regulated and sensitive industrial concerns to take advantage of the full spectrum of modern copyleft software, including code under the GNU General Public License, version 3 ("GPLv3"), and to manage their obligations under those licenses in ways that are commercially sound.

Software embedded in physical devices now determines how almost everything—from coffee pots and rice cookers to oil tankers and passenger airplanes—works. Safety and security, efficiency and repairability, fitness for purpose and adaptability to new conditions of all the physical products that we make and use now depends on our methods for developing, debugging, maintaining, securing and servicing the software embedded in them. These methods of "software governance," are to 21st-century technology what materials science and quality assurance practices were to 20th-century industrial activity. They are now a crucial "hidden input" to industry's ability to make, and government's ability to regulate, everything we use.

Few products encapsulate both the challenges and the possibilities in this area like the automobile. We stand on the verge of a transformative set of changes brought about by autopilot software capable of autonomously controlling some forms of "driving." But even now, the automobile is already a constellation of computers and an ecology of software. Even as currently designed for human operation and control, a contemporary car can include dozens of computers receiving input from hundreds of sensor devices and actuating everything from brakes to entertainment systems, all under the control of thousands of software components. Cars have long lifetimes, necessitating both scheduled and episodic maintenance. Safety testing and regulation of software involved in vehicle operation is extremely difficult, even given perfect trust between manufacturers and governments, which recent events have shown cannot be preserved. In addition, manufacturers have legitimate trade-secret and regulatory compliance interests which may conflict with the economic and social interests that

[*] Mark Shuttleworth is CEO of Canonical Ltd., which distributes the all-FOSS Ubuntu Core system.

[†] Eben Moglen is Professor of Law at Columbia Law School and Founding Director of the Software Freedom Law Center.



Automotive Software Governance and Copyleft 2

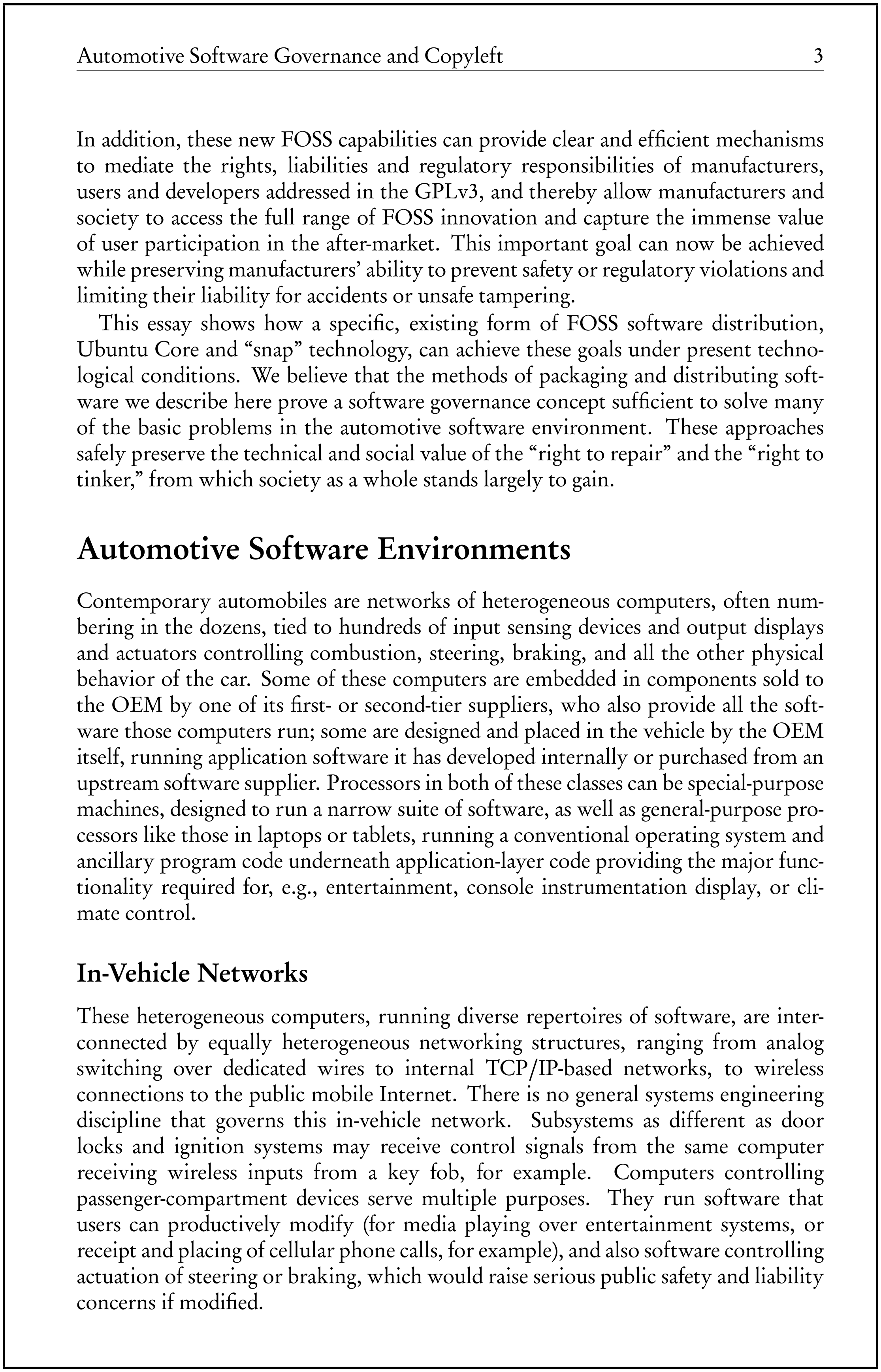
are served when the software in vehicles can be inspected, tested, and repaired by the largest number of qualified parties. Our 20th-century experience with automobiles showed the value for manufacturers in following the innovations created by car owners themselves, because people could learn to fix, adapt, and reuse automotive technologies by working “after-market” on cars.

For these reasons, society and the industry have much to gain from the use of a software governance system based on “free and open source” software (“FOSS”) in automobiles, such as Ubuntu Core—produced by Canonical Ltd.—which creates a framework that can be used to balance potentially competing objectives. FOSS is software produced and distributed under rules that give purchasers and users both the legal rights and technical enablements (like possession of source code and a means to install and use fixed or modified versions of programs) necessary to study, improve and share. FOSS has become the most important infrastructure material in software over the last generation. It includes the Android operating system in the vast majority of the world’s smartphones and tablets, and the GNU/Linux operating system dominant in server computers. It enables the utility computing power around the world we call “the cloud.” It comprises the software infrastructure on which the world’s digital “platforms” like Google, Facebook and Twitter run. FOSS is everywhere.

The pace of innovation in the automotive sector, as in other technology-centred sectors, has accelerated to the point where it is only economically feasible to compete by using shared and open code. FOSS has become central to the race to create better automotive experiences, and every established manufacturer and startup in the sector is dependent on a body of shared code under FOSS licenses. However, the automotive industry has tended to limit its own access to widespread innovation by avoiding FOSS under copyleft licenses which facilitate user modification, for lack of an effective means to manage those obligations in a commercially satisfactory manner.

In the automotive software environment, new technical capabilities in FOSS can help to solve the profound engineering and social problems of governance, security and liability from which we already suffer and which the rapid technological changes occurring in the industry will aggravate. These newly available capabilities and methods of software distribution can:

- Guarantee manufacturers’, owners’ and regulators’ precise knowledge of which version of what software is installed in any vehicle at any moment;
- Enable efficient, secure and reliable updates to be performed “over the air” or offline;
- Limit the access of individual software components with very precise granularity;
- Determine who has serviced or modified that software;
- Restore erroneously or incompletely modified software to a known, reliable state; and
- Even install or revert software upgrades and fixes during vehicle operation.



In addition, these new FOSS capabilities can provide clear and efficient mechanisms to mediate the rights, liabilities and regulatory responsibilities of manufacturers, users and developers addressed in the GPLv3, and thereby allow manufacturers and society to access the full range of FOSS innovation and capture the immense value of user participation in the after-market. This important goal can now be achieved while preserving manufacturers' ability to prevent safety or regulatory violations and limiting their liability for accidents or unsafe tampering.

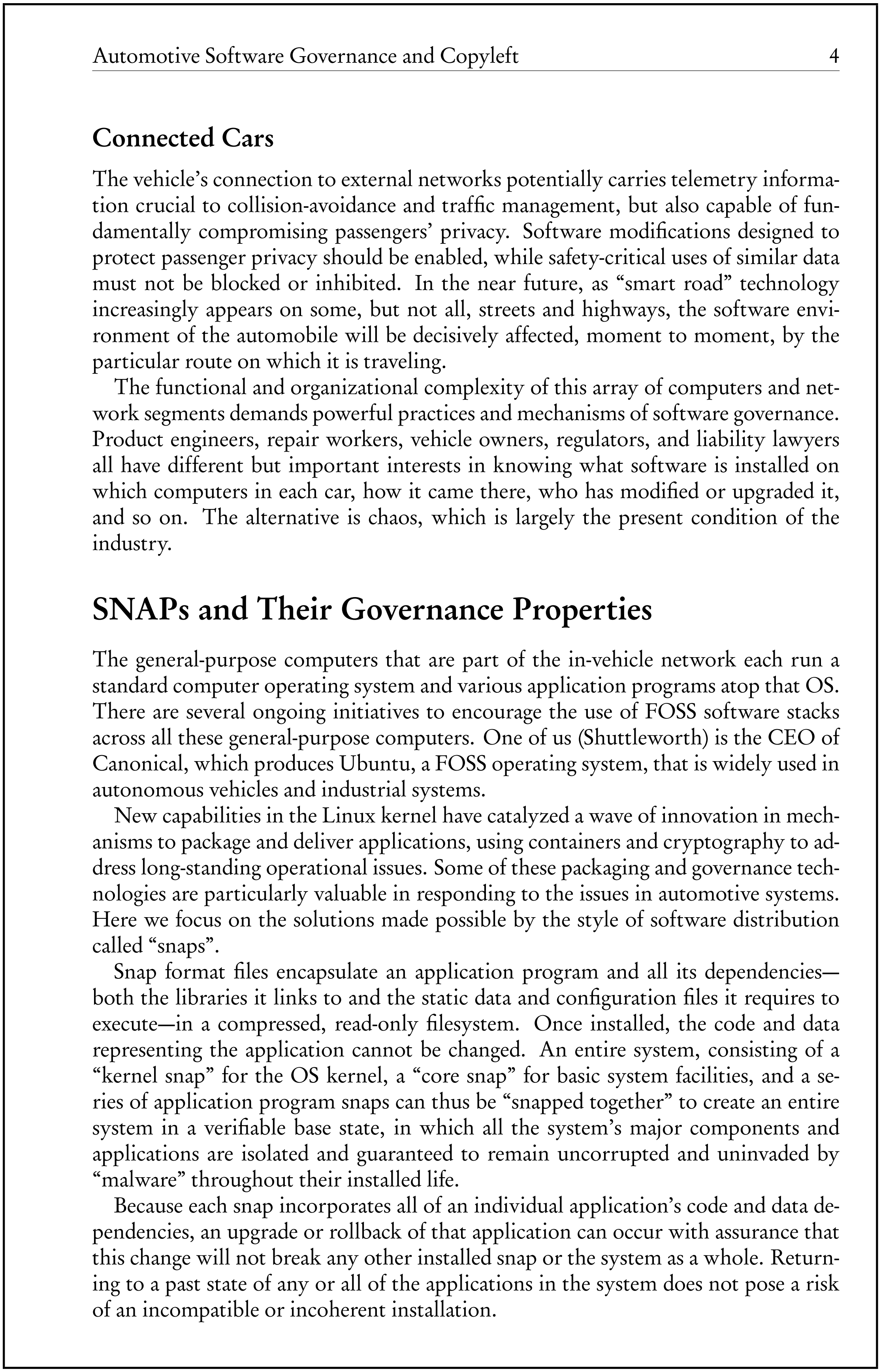
This essay shows how a specific, existing form of FOSS software distribution, Ubuntu Core and "snap" technology, can achieve these goals under present technological conditions. We believe that the methods of packaging and distributing software we describe here prove a software governance concept sufficient to solve many of the basic problems in the automotive software environment. These approaches safely preserve the technical and social value of the "right to repair" and the "right to tinker," from which society as a whole stands largely to gain.

Automotive Software Environments

Contemporary automobiles are networks of heterogeneous computers, often numbering in the dozens, tied to hundreds of input sensing devices and output displays and actuators controlling combustion, steering, braking, and all the other physical behavior of the car. Some of these computers are embedded in components sold to the OEM by one of its first- or second-tier suppliers, who also provide all the software those computers run; some are designed and placed in the vehicle by the OEM itself, running application software it has developed internally or purchased from an upstream software supplier. Processors in both of these classes can be special-purpose machines, designed to run a narrow suite of software, as well as general-purpose processors like those in laptops or tablets, running a conventional operating system and ancillary program code underneath application-layer code providing the major functionality required for, e.g., entertainment, console instrumentation display, or climate control.

In-Vehicle Networks

These heterogeneous computers, running diverse repertoires of software, are interconnected by equally heterogeneous networking structures, ranging from analog switching over dedicated wires to internal TCP/IP-based networks, to wireless connections to the public mobile Internet. There is no general systems engineering discipline that governs this in-vehicle network. Subsystems as different as door locks and ignition systems may receive control signals from the same computer receiving wireless inputs from a key fob, for example. Computers controlling passenger-compartment devices serve multiple purposes. They run software that users can productively modify (for media playing over entertainment systems, or receipt and placing of cellular phone calls, for example), and also software controlling actuation of steering or braking, which would raise serious public safety and liability concerns if modified.



Connected Cars

The vehicle's connection to external networks potentially carries telemetry information crucial to collision-avoidance and traffic management, but also capable of fundamentally compromising passengers' privacy. Software modifications designed to protect passenger privacy should be enabled, while safety-critical uses of similar data must not be blocked or inhibited. In the near future, as "smart road" technology increasingly appears on some, but not all, streets and highways, the software environment of the automobile will be decisively affected, moment to moment, by the particular route on which it is traveling.

The functional and organizational complexity of this array of computers and network segments demands powerful practices and mechanisms of software governance. Product engineers, repair workers, vehicle owners, regulators, and liability lawyers all have different but important interests in knowing what software is installed on which computers in each car, how it came there, who has modified or upgraded it, and so on. The alternative is chaos, which is largely the present condition of the industry.

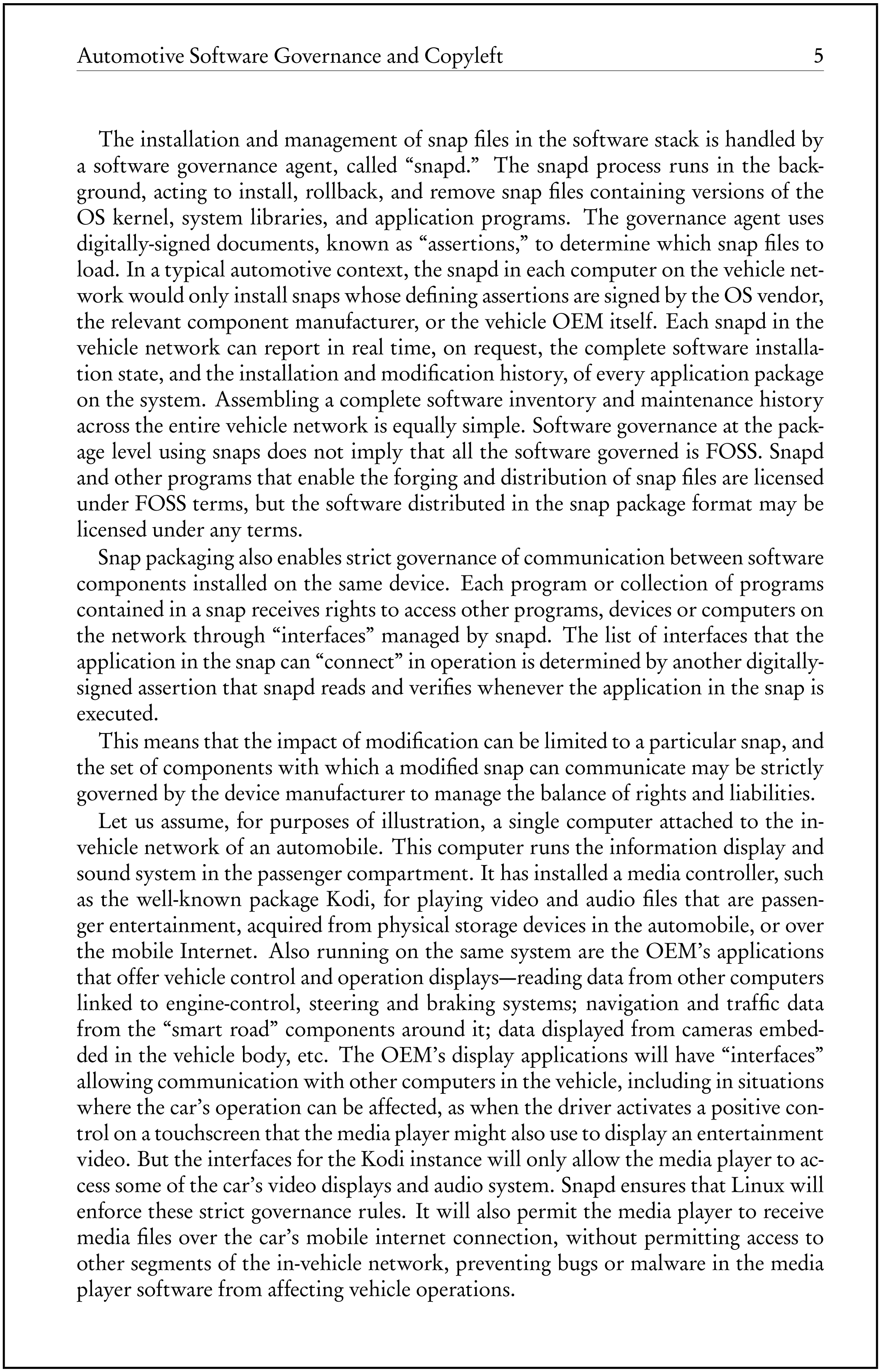
SNAPs and Their Governance Properties

The general-purpose computers that are part of the in-vehicle network each run a standard computer operating system and various application programs atop that OS. There are several ongoing initiatives to encourage the use of FOSS software stacks across all these general-purpose computers. One of us (Shuttleworth) is the CEO of Canonical, which produces Ubuntu, a FOSS operating system, that is widely used in autonomous vehicles and industrial systems.

New capabilities in the Linux kernel have catalyzed a wave of innovation in mechanisms to package and deliver applications, using containers and cryptography to address long-standing operational issues. Some of these packaging and governance technologies are particularly valuable in responding to the issues in automotive systems. Here we focus on the solutions made possible by the style of software distribution called "snaps".

Snap format files encapsulate an application program and all its dependencies—both the libraries it links to and the static data and configuration files it requires to execute—in a compressed, read-only filesystem. Once installed, the code and data representing the application cannot be changed. An entire system, consisting of a "kernel snap" for the OS kernel, a "core snap" for basic system facilities, and a series of application program snaps can thus be "snapped together" to create an entire system in a verifiable base state, in which all the system's major components and applications are isolated and guaranteed to remain uncorrupted and uninvaded by "malware" throughout their installed life.

Because each snap incorporates all of an individual application's code and data dependencies, an upgrade or rollback of that application can occur with assurance that this change will not break any other installed snap or the system as a whole. Returning to a past state of any or all of the applications in the system does not pose a risk of an incompatible or incoherent installation.

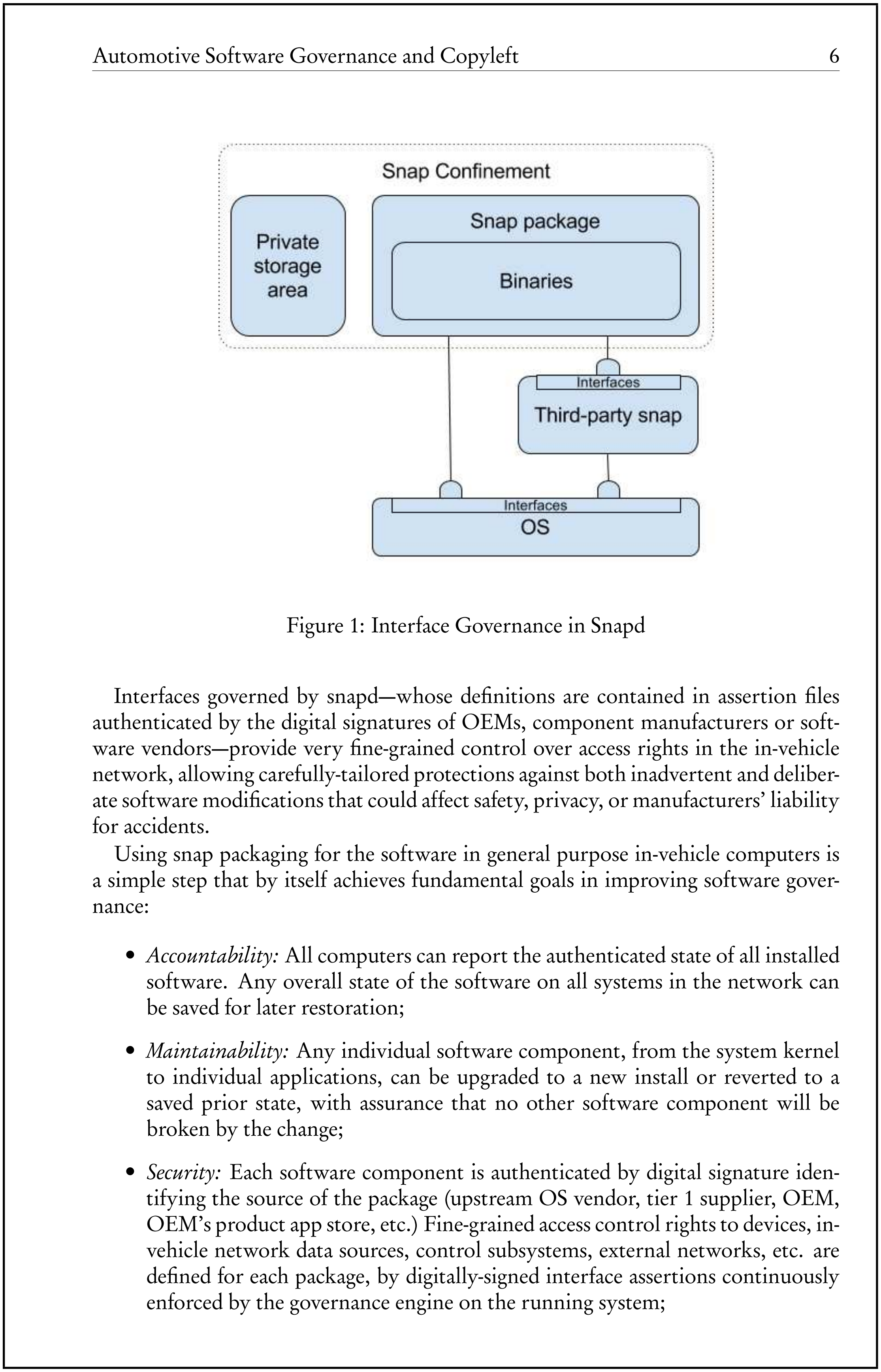


The installation and management of snap files in the software stack is handled by a software governance agent, called “snapd.” The snapd process runs in the background, acting to install, rollback, and remove snap files containing versions of the OS kernel, system libraries, and application programs. The governance agent uses digitally-signed documents, known as “assertions,” to determine which snap files to load. In a typical automotive context, the snapd in each computer on the vehicle network would only install snaps whose defining assertions are signed by the OS vendor, the relevant component manufacturer, or the vehicle OEM itself. Each snapd in the vehicle network can report in real time, on request, the complete software installation state, and the installation and modification history, of every application package on the system. Assembling a complete software inventory and maintenance history across the entire vehicle network is equally simple. Software governance at the package level using snaps does not imply that all the software governed is FOSS. Snapd and other programs that enable the forging and distribution of snap files are licensed under FOSS terms, but the software distributed in the snap package format may be licensed under any terms.

Snap packaging also enables strict governance of communication between software components installed on the same device. Each program or collection of programs contained in a snap receives rights to access other programs, devices or computers on the network through “interfaces” managed by snapd. The list of interfaces that the application in the snap can “connect” in operation is determined by another digitally-signed assertion that snapd reads and verifies whenever the application in the snap is executed.

This means that the impact of modification can be limited to a particular snap, and the set of components with which a modified snap can communicate may be strictly governed by the device manufacturer to manage the balance of rights and liabilities.

Let us assume, for purposes of illustration, a single computer attached to the in-vehicle network of an automobile. This computer runs the information display and sound system in the passenger compartment. It has installed a media controller, such as the well-known package Kodi, for playing video and audio files that are passenger entertainment, acquired from physical storage devices in the automobile, or over the mobile Internet. Also running on the same system are the OEM’s applications that offer vehicle control and operation displays—reading data from other computers linked to engine-control, steering and braking systems; navigation and traffic data from the “smart road” components around it; data displayed from cameras embedded in the vehicle body, etc. The OEM’s display applications will have “interfaces” allowing communication with other computers in the vehicle, including in situations where the car’s operation can be affected, as when the driver activates a positive control on a touchscreen that the media player might also use to display an entertainment video. But the interfaces for the Kodi instance will only allow the media player to access some of the car’s video displays and audio system. Snapd ensures that Linux will enforce these strict governance rules. It will also permit the media player to receive media files over the car’s mobile internet connection, without permitting access to other segments of the in-vehicle network, preventing bugs or malware in the media player software from affecting vehicle operations.



Automotive Software Governance and Copyleft 6

Figure 1: Interface Governance in Snapd

Interfaces governed by snapd—whose definitions are contained in assertion files authenticated by the digital signatures of OEMs, component manufacturers or software vendors—provide very fine-grained control over access rights in the in-vehicle network, allowing carefully-tailored protections against both inadvertent and deliberate software modifications that could affect safety, privacy, or manufacturers' liability for accidents.

Using snap packaging for the software in general purpose in-vehicle computers is a simple step that by itself achieves fundamental goals in improving software governance:

- *Accountability:* All computers can report the authenticated state of all installed software. Any overall state of the software on all systems in the network can be saved for later restoration;

- *Maintainability:* Any individual software component, from the system kernel to individual applications, can be upgraded to a new install or reverted to a saved prior state, with assurance that no other software component will be broken by the change;

- *Security:* Each software component is authenticated by digital signature identifying the source of the package (upstream OS vendor, tier 1 supplier, OEM, OEM's product app store, etc.) Fine-grained access control rights to devices, in-vehicle network data sources, control subsystems, external networks, etc. are defined for each package, by digitally-signed interface assertions continuously enforced by the governance engine on the running system;

- *Adaptability:* Multiple states of each software package throughout the vehicle, and overall states representing a snapshot of all the software on all systems, can be saved and restored with assurance that each version of each package will operate correctly. All such state changes are auditable for maintenance and other purposes. The vehicle's software environment therefore gains comprehensive adaptability: it can be changed both globally and locally to respond to operating conditions, to deal with hardware failures, and to capture user innovation.

Copyleft and the Right to Tinker

The adaptability in software governance brought about by this simple change in software packaging bears directly on the ability to capture user innovation in vehicle software, because it also allows us to govern user-modified and other experimental versions of FOSS software effectively.

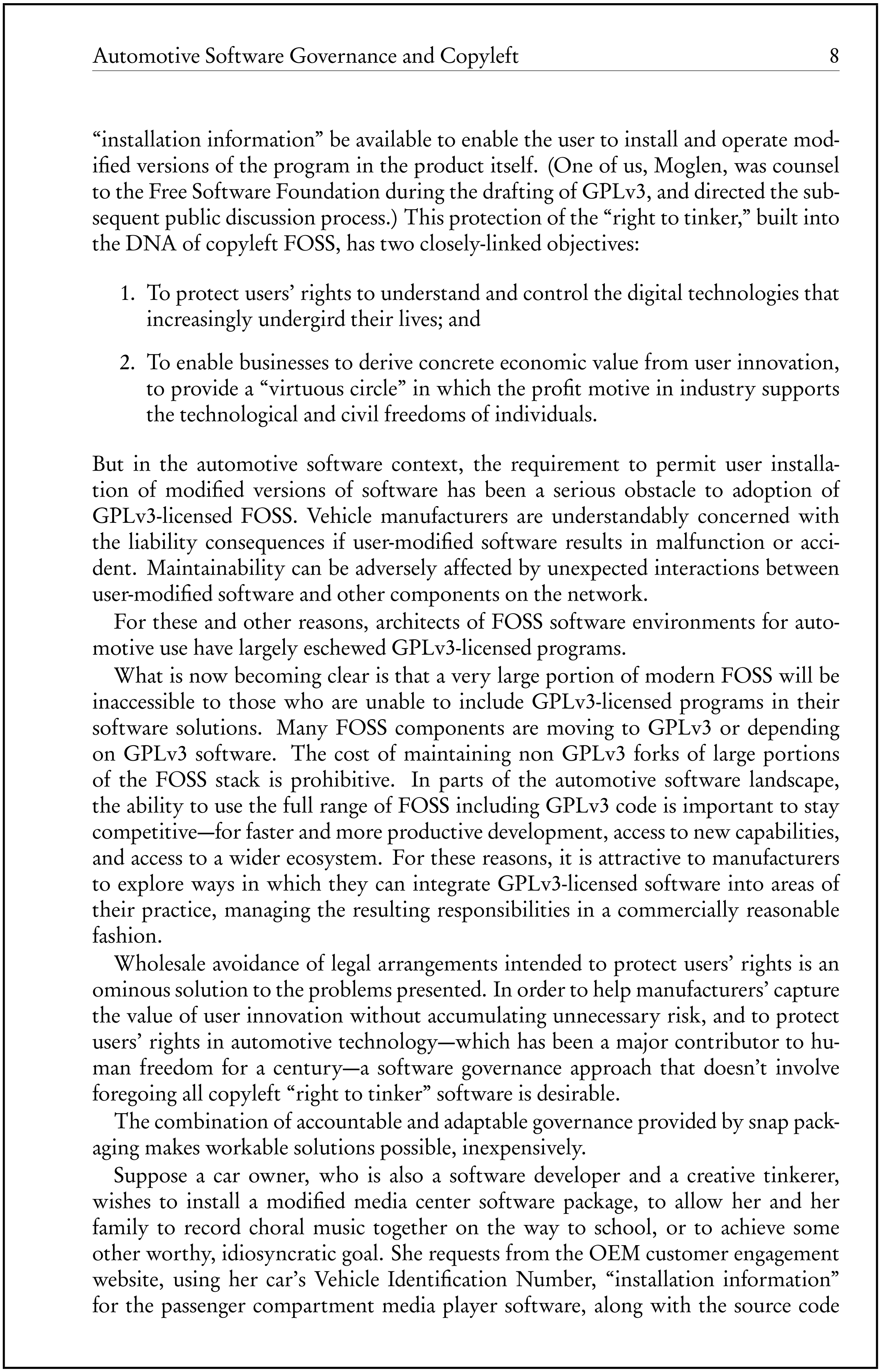
Capturing user innovation can only occur if we enable car owners to modify the vehicle's software environment. This openness to downstream innovation is the heart of the success of FOSS as a model for making software. The history of automotive technology also shows the enormous importance of user ingenuity in seeding and stimulating technological advancement.

Preserving the "right to tinker" yields larger social gains, beyond the economic value to manufacturers of user innovation. The free after-market in repair services, which includes software maintenance along with other forms of vehicle service, is a source of employment and small business activity. As ride-sharing and fleet services, such as Lyft, Uber and Zipcar, seek to limit individual ownership and operation of self-driving cars, in part on the ground that other entities lack the expertise to perform such complex software maintenance,[1] it becomes apparent that the subject also involves basic competition law issues.

The copyright licenses that make FOSS possible can be divided into two categories. So-called "permissive" licenses allow programmers to place under any license of their choice programs made in part from FOSS parts. This licensing structure maximizes the choices available to software developers. The category of "copyleft" licenses, on the other hand, requires programmers and firms using copylefted FOSS parts to release their own work under the same or equivalent copyleft licenses. This licensing structure emphasizes the rights of *users* to receive program source code and the other materials necessary to enable experimentation, improvement, and adaptation to new purposes. The most widely-used and influential copyleft licenses are in the family of the GNU General Public License, published by the Free Software Foundation and applied not only to its GNU operating system, but to hundreds of thousands of other computer programs used in all technical and social contexts.

GPLv3, a license promulgated in 2007 after lengthy public review and discussion, requires not only that GPLv3 programs and all works based on them enable users to modify and share their modified versions, but also requires that when such programs are distributed embedded in a "user product" (which includes an automobile) that

[1] See, for example, www.sharedmobilityprinciples.org, point 10.



Automotive Software Governance and Copyleft 8

"installation information" be available to enable the user to install and operate modified versions of the program in the product itself. (One of us, Moglen, was counsel to the Free Software Foundation during the drafting of GPLv3, and directed the subsequent public discussion process.) This protection of the "right to tinker," built into the DNA of copyleft FOSS, has two closely-linked objectives:

1. To protect users' rights to understand and control the digital technologies that increasingly undergird their lives; and

2. To enable businesses to derive concrete economic value from user innovation, to provide a "virtuous circle" in which the profit motive in industry supports the technological and civil freedoms of individuals.

But in the automotive software context, the requirement to permit user installation of modified versions of software has been a serious obstacle to adoption of GPLv3-licensed FOSS. Vehicle manufacturers are understandably concerned with the liability consequences if user-modified software results in malfunction or accident. Maintainability can be adversely affected by unexpected interactions between user-modified software and other components on the network.

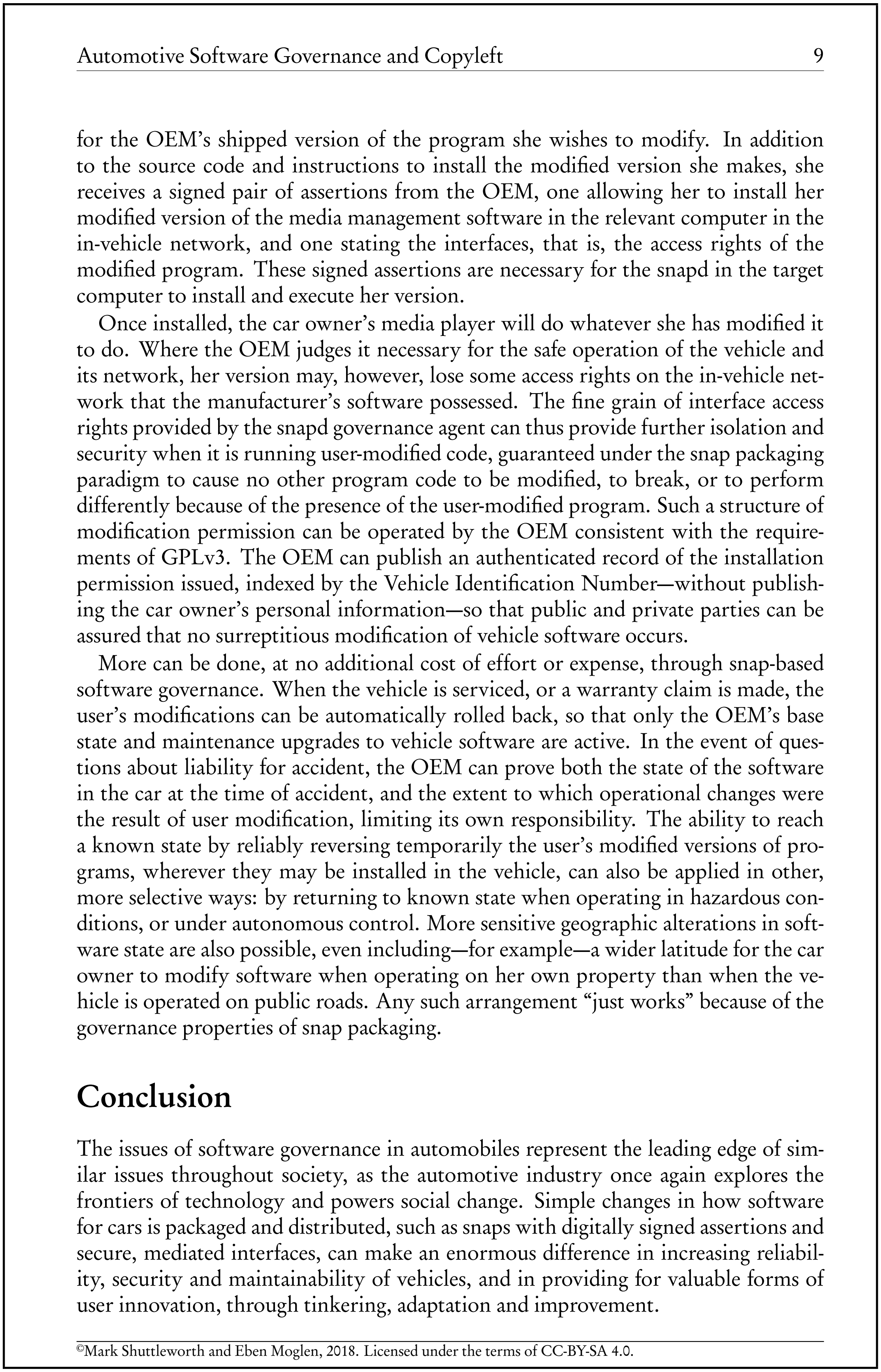
For these and other reasons, architects of FOSS software environments for automotive use have largely eschewed GPLv3-licensed programs.

What is now becoming clear is that a very large portion of modern FOSS will be inaccessible to those who are unable to include GPLv3-licensed programs in their software solutions. Many FOSS components are moving to GPLv3 or depending on GPLv3 software. The cost of maintaining non GPLv3 forks of large portions of the FOSS stack is prohibitive. In parts of the automotive software landscape, the ability to use the full range of FOSS including GPLv3 code is important to stay competitive—for faster and more productive development, access to new capabilities, and access to a wider ecosystem. For these reasons, it is attractive to manufacturers to explore ways in which they can integrate GPLv3-licensed software into areas of their practice, managing the resulting responsibilities in a commercially reasonable fashion.

Wholesale avoidance of legal arrangements intended to protect users' rights is an ominous solution to the problems presented. In order to help manufacturers' capture the value of user innovation without accumulating unnecessary risk, and to protect users' rights in automotive technology—which has been a major contributor to human freedom for a century—a software governance approach that doesn't involve foregoing all copyleft "right to tinker" software is desirable.

The combination of accountable and adaptable governance provided by snap packaging makes workable solutions possible, inexpensively.

Suppose a car owner, who is also a software developer and a creative tinkerer, wishes to install a modified media center software package, to allow her and her family to record choral music together on the way to school, or to achieve some other worthy, idiosyncratic goal. She requests from the OEM customer engagement website, using her car's Vehicle Identification Number, "installation information" for the passenger compartment media player software, along with the source code



for the OEM's shipped version of the program she wishes to modify. In addition to the source code and instructions to install the modified version she makes, she receives a signed pair of assertions from the OEM, one allowing her to install her modified version of the media management software in the relevant computer in the in-vehicle network, and one stating the interfaces, that is, the access rights of the modified program. These signed assertions are necessary for the snapd in the target computer to install and execute her version.

Once installed, the car owner's media player will do whatever she has modified it to do. Where the OEM judges it necessary for the safe operation of the vehicle and its network, her version may, however, lose some access rights on the in-vehicle network that the manufacturer's software possessed. The fine grain of interface access rights provided by the snapd governance agent can thus provide further isolation and security when it is running user-modified code, guaranteed under the snap packaging paradigm to cause no other program code to be modified, to break, or to perform differently because of the presence of the user-modified program. Such a structure of modification permission can be operated by the OEM consistent with the requirements of GPLv3. The OEM can publish an authenticated record of the installation permission issued, indexed by the Vehicle Identification Number—without publishing the car owner's personal information—so that public and private parties can be assured that no surreptitious modification of vehicle software occurs.

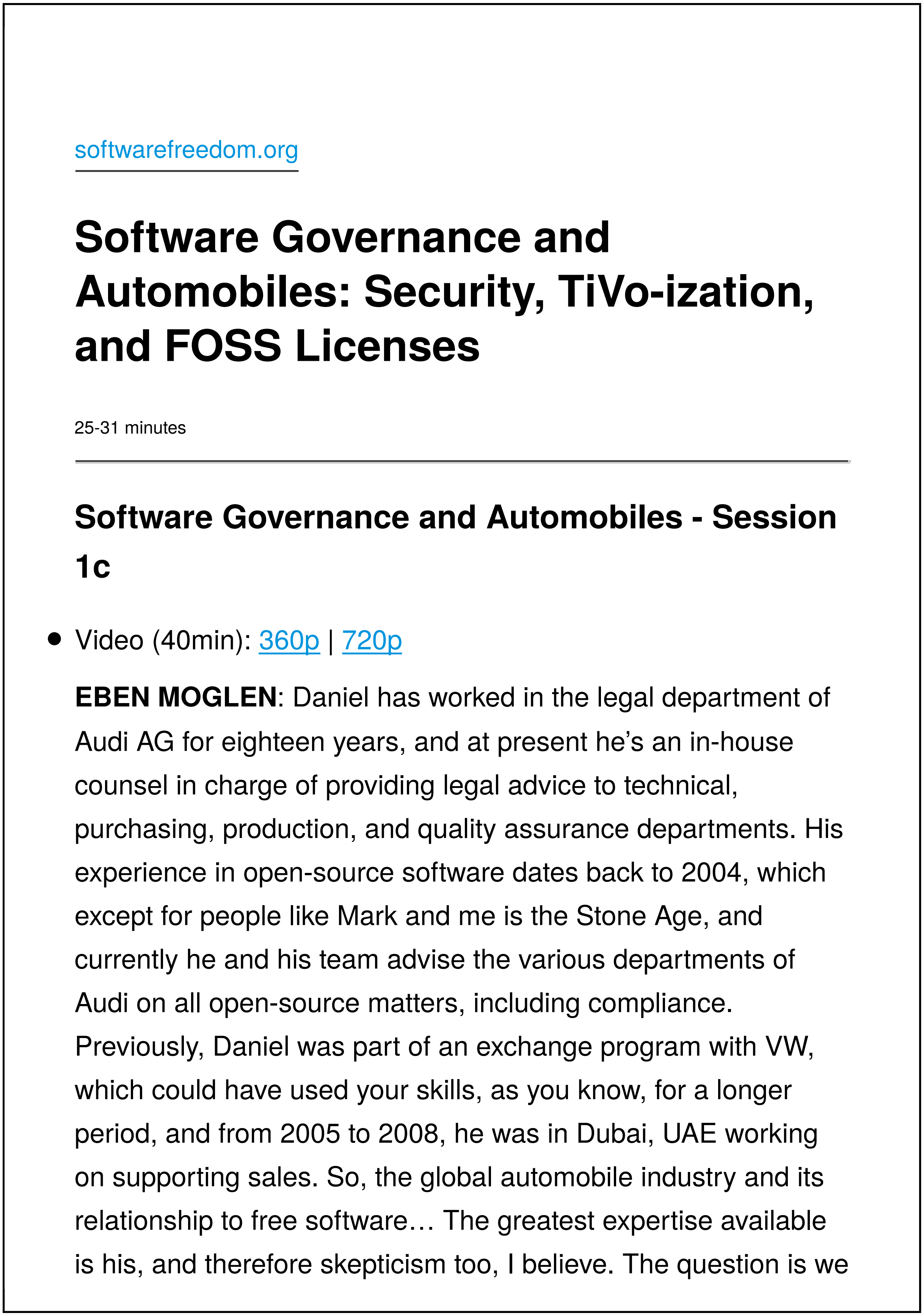
More can be done, at no additional cost of effort or expense, through snap-based software governance. When the vehicle is serviced, or a warranty claim is made, the user's modifications can be automatically rolled back, so that only the OEM's base state and maintenance upgrades to vehicle software are active. In the event of questions about liability for accident, the OEM can prove both the state of the software in the car at the time of accident, and the extent to which operational changes were the result of user modification, limiting its own responsibility. The ability to reach a known state by reliably reversing temporarily the user's modified versions of programs, wherever they may be installed in the vehicle, can also be applied in other, more selective ways: by returning to known state when operating in hazardous conditions, or under autonomous control. More sensitive geographic alterations in software state are also possible, even including—for example—a wider latitude for the car owner to modify software when operating on her own property than when the vehicle is operated on public roads. Any such arrangement "just works" because of the governance properties of snap packaging.

Conclusion

The issues of software governance in automobiles represent the leading edge of similar issues throughout society, as the automotive industry once again explores the frontiers of technology and powers social change. Simple changes in how software for cars is packaged and distributed, such as snaps with digitally signed assertions and secure, mediated interfaces, can make an enormous difference in increasing reliability, security and maintainability of vehicles, and in providing for valuable forms of user innovation, through tinkering, adaptation and improvement.

Chapter 15

Security, TiVo-ization, and FOSS Licenses (Daniel Patnaik)



softwarefreedom.org

Software Governance and Automobiles: Security, TiVo-ization, and FOSS Licenses

25-31 minutes

Software Governance and Automobiles - Session 1c

- Video (40min): 360p | 720p

EBEN MOGLEN: Daniel has worked in the legal department of Audi AG for eighteen years, and at present he’s an in-house counsel in charge of providing legal advice to technical, purchasing, production, and quality assurance departments. His experience in open-source software dates back to 2004, which except for people like Mark and me is the Stone Age, and currently he and his team advise the various departments of Audi on all open-source matters, including compliance. Previously, Daniel was part of an exchange program with VW, which could have used your skills, as you know, for a longer period, and from 2005 to 2008, he was in Dubai, UAE working on supporting sales. So, the global automobile industry and its relationship to free software… The greatest expertise available is his, and therefore skepticism too, I believe. The question is we have these ideas, now it's time for the automobile industry to kick the tires. Would you mind doing so for me, please? Thank you. Please welcome Daniel Patnaik.

DANIEL PATNAIK: Perfect. Thank you very much, ladies and gentlemen, hello–thank you for having me here. First of all, I would like to start with an excuse. Please excuse my English. My English is not as good as Eben's and maybe Mark's English. I am a native German lawyer, and so forgive me if I'm lacking some English words. I'm trying to do my best. I also have to state that whatever I am going to present today is a statement of me personally. It's not an official statement of my employer, Audi AG, but, however, you can and will see how I understand things and, of course, how we in the automotive industry see some of the points which have been raised here by Eben and by Mark. I think it's a very, very interesting topic, and we are deeply working on it, but there are a lot of challenges, which were already highlighted by both of you, which is absolutely right. We are in a very complex matter and very complex area, and I would like to give you an industry view, a view of how that can be seen from the automotive industry perspective.

For all of you who are maybe not so familiar with TiVo and the TiVo case, TiVo, from my point of view, is a term for a security measure which is in due course already widely used in commercial products, not only the automotive industry but any other industry, which shall prevent a special software which is maybe signed or not signed from running on a specific system.

So, that means the user cannot–he may change software–but he cannot install such a modified version of a software component. This, however, and this is, in general, in a kind of a conflict is, however, required by open source license.

In this case, as I have put up here, the hardware checks of a software for an expected signature are there and if the system notices that there is a change it shuts down, if it finds out that there's not a match. So, whatever has been checked before isn't there or has been modified. This can be done by using build-in routines, which may require those checks in a hardware.

The term TiVo-ization, as I have it up here, is derived from a specific case–the TiVo case. TiVo and Linux… I think the general topic there was a digital video receiver. It contained Linux, which was under the GPLv2 license. The source code for the system was available but it contained a technical blocking device which prevented users from running any modified version of a software, and the Free Software Foundation took up the case, complained that the users at the end were prevented from using or from exercising their freedom to change, to alter and to exercise, at least, their right to do changes and operate it in the system.

The third aspect that I want to touch here is, yes, that was the case, but in the automotive industry I think we have security and safety issues. We want to prevent that we have, kind of, manipulation. Why is that important to us? It's important to us because we want to safeguard our products and, therefore, we have, of course, signature checks through Secure Boot, Chain of Trust–and this is widely used–without a possibility to release signature checks.

And I think the topic in the later part of the day today, which talks about autonomous cars–and this is what I've put up here as a small picture, illustration–this is a little bit of the trend, and I think Mark and Eben have been touching on that as well. This is the trend we see. Imagine autonomous cars with a software modified by users. So, this is just a general, a very flashlight, here.

You can think now that autonomous cars are coming to the cities and are being tested. We are reading a lot about what's happening. Now, also, the countries, the cities, and the states, are trying to regulate and try to get permissions–who can use it and who can do it–but now imagine not only the car manufacturer will do a certain software algorithm to do that kind of, to produce that kind of software but also maybe a customer, a user, will work on that and then ultimately run his car on that kind of software and at the end of the day, the car is on the roads and some difficulties might occur, which then start from personal injury to even more dangerous issues.

So this is the trend in a very highly abstract way that we see, but I also want to go and dive more into the deep what that means at the end of the day. And, of course, if you have questions, please always raise your hands and we can ask you.

I put up this nice, very uncomplicated, slide here. It looks very highly dense with information, but I think I'm just going to explain to you what it means, and, of course, it starts with open source software. We try to cluster software a little bit in that way. Of course, everyone can do the clustering in their own way. This is the way that I see the things–or we see the things–and, but of course, there are different ways to cluster software and to look at it.

So, we have said open source software can be differentiated in strong Copyleft, limited Copyleft, and without Copyleft, and I have listed the different, the major, and the most important open source licenses here from GPLv3 to GPLv2 up to BSD and Apache.

Now, if you look at all of those open source licenses, we can look at the… Maybe I'll start with the green part, and whatever is green here–whatever is in the green box–has a explanation–I shouldn't walk away too far from my microphone, so you can all hear me.

So, if you look at the green part, we can find out, and if you look at the licenses, you'll find out that most licenses which are marked with a green frame, that those licenses do not collide with the TiVo case or TiVo-ization. There are no specific TiVo-ization clauses, wording, which is seen in GPL or other non-Copyleft licenses.

Implementation of the system with hardware, and, therefore, signature checks is possible. So, I think this is also a very important message that there is already software out there which allows both things. First of all, signature checks but also using open source software. So, that's an important point, and if you see very up there in the high part is GPLv2, which, from my perspective, allows signature checks but, nevertheless, using open source software.

Now, if you come to the yellow part, which is GPLv2.1 or

GPLv3, for example, with the runtime exception. Here we can see that we might have a diversity of interpretations for some of the open source licenses. There is no specific, if you look at the GPLv2.1, there's no specific provision in the license which relates to the TiVo-ization of a software. But, as I have stated it up here, the wording relates and refers to exchange Exchange means that the customer needs to be in a position to exchange, to alter, to change, the code.

The exception provision here is liberating from the strong requirements of GPL. You could possibly understand that this could be interpreted in the way that includes the TiVo-ization issue, but that's all part of interpretation. It's not a very clear understanding, at least from my side, what that means at the end. But there's a possibility of interpretation towards TiVo-ization.

And then the top part, which are the red licenses, which I have listed here, including GPLv3, where it's clearly stated that TiVo-ization is prohibited in any new version of the license, which is, as I have stated here, it's GPLv3, it's GPLv3 or LGPLv3, which clearly and verbally interdicts TiVo-ization. A license-compliant implementation of a TiVo-ized system is only possible, from my understanding, if we have a release of the possible signature keys.

So, this is a way of how you can cluster the open source–free and open source–software licenses, and you can have a view on what that means on how this relates to TiVo-ization.

Often discussed–that's how I start my next slide–is LGPLv2.1, where you have seen if you remember from the page before, it’s in the yellow corner. What does it mean if you talk about and if you see and look towards a exchangeability and the TiVo-ization issue? What does–and, maybe, it also relates a little bit to other licenses as well, LGPLv2.1 states that the system shall operate properly if the user exchanges and modifies the LGPL components. I’ve put down, also, part of the license–I don’t know if you’re able to read it, it’s a little small, but it’s highlighted in the yellow part here, and I can read it out so you can read it and you can understand it: so that the user can modify the library and then re-link to produce a modified executable containing the modified library or, which is then (6.B), will operate properly with a modified version of the library if the user installs one.

So, TiVo-ization, if you make a security check, at the end of the day prevents an exchange of software components so that the user is not able to exchange LGPLv2.1 components in a TiVo-ized system and gets a running system, yet, at least from my point of view, I have not seen a specific clause interdicting TiVo-ization as it is–as I have said and mentioned before on the slide before–was inserted for version 3.1 of the GPL. There’s no specific jurisdiction or law cases on the specific topic up to now, so, from my point of view, it’s a question of interpretation of the license text: if GPLv2.1 interdicts TiVo-ization or not.

So, what are the different ways of interpreting that license or the exchangeability of the license? I have listed here four ways of interpretation. You can have a very panicked interpretation and say, “Okay, we have to avoid…”–and this is up to a company policy, of course, well, then, your company's or your organization's. You can have a very panicked interpretation and say, "Okay we have to avoid all possible risk and interdict the use of open source software, Copyleft components, LGPL, or even GPL entirely." That means you are not going to implement GNU and Linux-based operating systems, possibly, as LGPLv2.1 or, even, GPLv3 components are included. I put it here, is that maybe outdated? That's what also Eben and Mark have talked about. Linux is widely used in industry, so, as you said, if you're too strict on that you might lose the path to innovation. So, that's absolutely the right point. That's why I put it here as a call-out.

You can also have a–this is maybe the next level–a conservative interpretation. There is a potential legal risk for non-compliance because of the way of interpretation of that clause. You can avoid implementation of such a software in your TiVo-ized systems or in technical construction for… Or, you can just avoid using it entirely or you can use technical construction for exchangeability. I think these are ways to cover this if you have a conservative interpretation.

Or you can go to the liberal interpretation: you can say there's no explicit wording or case-law interdicting TiVo-ization, so you can say that implementation is possible, especially if it's unavoidable from a technical point of view. So, these are two ways in the middle.

And you can say to have an indifferent interpretation: you can say that up to now most of the legal claims based on a Copyleft issue were not touching TiVo-ization, so the compliance with

Copyleft was the main goal. So you can say, “Okay, I’m indifferent on that”, so what I put up here, it might be risky if you look at the original case, Linux v. TiVo."

So, I think you have two very extreme positions, the panicked interpretation and the very indifferent interpretation where they say, “Who cares?” I think these are not the right ways to look at it. Therefore, and that’s why I’ve framed the both positions here, it’s probably more going towards positions two and three, and you have to look at the legal risks and how you may find a way to cope with it.

The area of conflict which we have been touching already this morning, with regard to TiVo-ization, is that–and this is probably not only applying to the automotive industry–is that developers want to secure a software against manipulation. As Eben also mentioned, you want to protect your intellectual property and implement certain technical features without showing everything to your competitors. That’s important for us.

But, on another point, which is also important, there’s a legal need for compliance with implemented open source components and the underlying software licenses because of potential claims. There might be a claim for damage, a claim for callback–because you have to bring back your product in order to rectify, to exchange, to bring it into license compliance, and for a product which is sold around the world, which is distributed heavily, this can be quite a high danger. It could be very cost-imminent, and the cost could be very high.

Sometimes, and this is the third point, a liberal or conservative interpretation regarding the license provisions is possible, so you have to really look at the specific case. So, the safest option, the legally safest option, if someone says, “OK, Daniel, what is your interpretation or what is your recommendation if I take a very safe option?” And I would say, “Yes, do not use that if we have TiVo-ized system–so, a secured system.” This could be really the legally safest option, but this is really the question, whether this brings us forward, as Mark and Eben rightly mentioned.

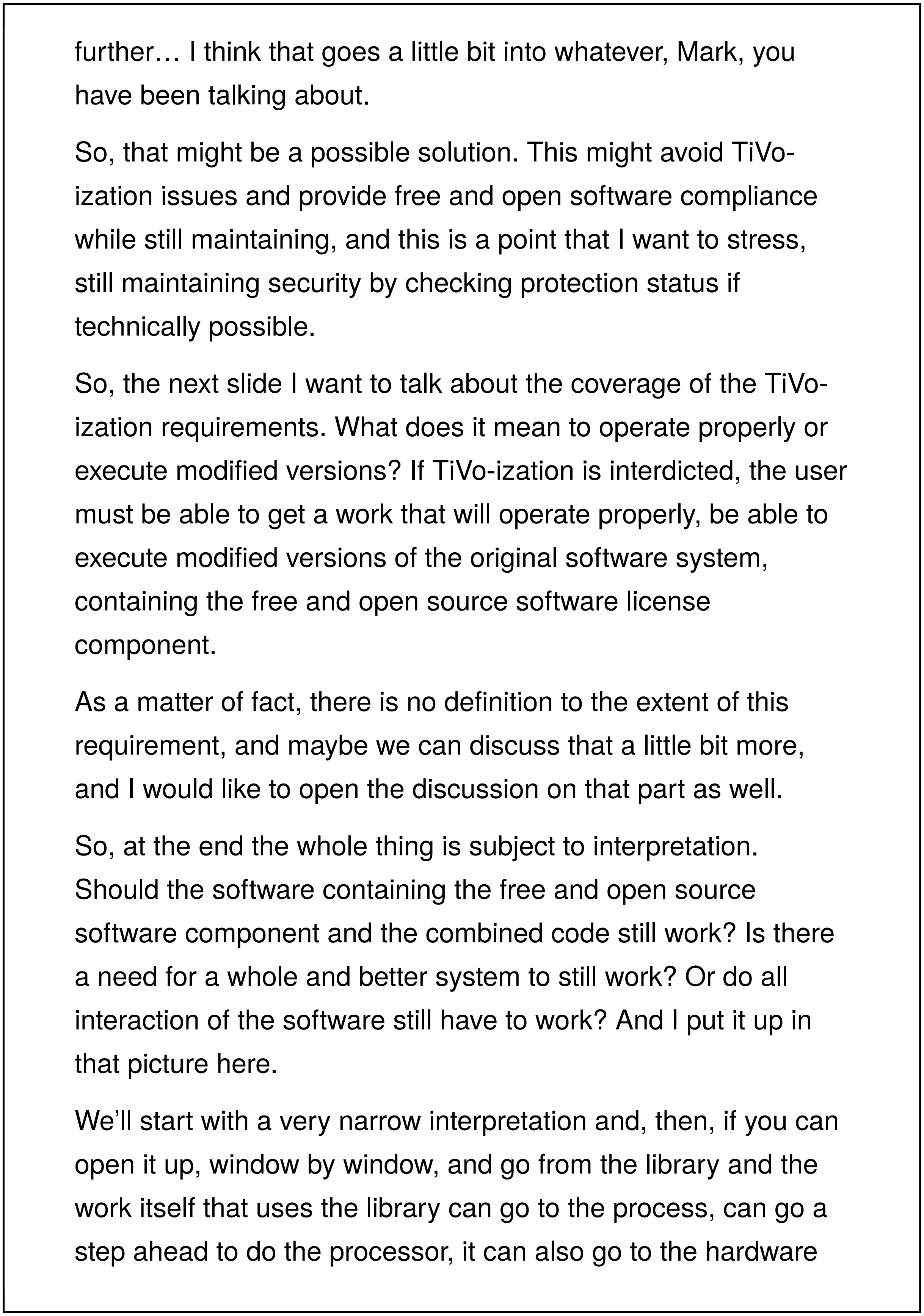
You can also say, “I can make sure the user is able to exchange the respective open-source component with modified versions of the library and still get a work that will operate properly and to be able to execute modified versions.” How can you do that? You can put libraries, and I think it goes a little bit into what Mark and Eben pointed out, you can put libraries in a separate file system and include them from the signature checks, Secure Boot, Chain of Trust, or else you can give the user the possibility to obtain the signature keys–this is also another possibility. You can offer your code as an object code so the user can make a separate but working version of the software that is then at the end, not limited by the delivered hardware.

So, there are ways to do that. And that is all an issue which relates to your understanding of software compliance. So I mentioned here the four aspects, at least from my point of view as I mentioned it already: so, you have the security of the system, you have a technical need for a component, you have the protection of intellectual property or know-how, and you have the potential interpretation of licenses. Those are the four components which ring around open-source compliance, and you have to find your way of interpreting a way forward.

So, on the next slide, I put up an attempt for a solution for combining TiVo-ization with free and open source software compliance. So, on the left part, which I put in a red box, I showed a system: for example, you have a partition 1, which represents a file-system combining proprietary and public components–a library. You have an observer which is monitoring a protection by signature, as I mentioned it before, during the runtime–for safety and security reasons. In case of a mismatch, that's exactly what we're talking about, the observer may block the access to and the interaction with the library if protection for partition 1 is activated.

What does it mean as a result? This means a non-compliance in the case of a FOSS license with interdicting of TiVo-ization, as the system will not work with modified versions of the lib file or the library. So at the end of the day, I want to show, this is not possible here if you want to use it in a way as I've put it up here in the red box.

How could it maybe work or how can it work? You have a component, as I have mentioned, "lib," which is placed in a separate file system partition instead of inclusion, as in the red box, including it into partition 1 where it is needed. You can protect the partition lib by deactivating, on request of the user, allowing the exchange, and even if you look at the protection for partition lib and this is deactivated, the rest of the system, including partition 1 continues to be protected by the observer. I think, and I don't know if we have to discuss that a little bit



further… I think that goes a little bit into whatever, Mark, you have been talking about.

So, that might be a possible solution. This might avoid TiVo-ization issues and provide free and open software compliance while still maintaining, and this is a point that I want to stress, still maintaining security by checking protection status if technically possible.

So, the next slide I want to talk about the coverage of the TiVo-ization requirements. What does it mean to operate properly or execute modified versions? If TiVo-ization is interdicted, the user must be able to get a work that will operate properly, be able to execute modified versions of the original software system, containing the free and open source software license component.

As a matter of fact, there is no definition to the extent of this requirement, and maybe we can discuss that a little bit more, and I would like to open the discussion on that part as well.

So, at the end the whole thing is subject to interpretation. Should the software containing the free and open source software component and the combined code still work? Is there a need for a whole and better system to still work? Or do all interaction of the software still have to work? And I put it up in that picture here.

We'll start with a very narrow interpretation and, then, if you can open it up, window by window, and go from the library and the work itself that uses the library can go to the process, can go a step ahead to do the processor, it can also go to the hardware unit, it can go to the delivered systems–the computer network in an office–but you can also go to the, probably most and wide interpretation, or set of picture, which is the system and the external services.

So, where is the frame? Where is the area where you look at it? Is it… If you look at the smallest interpretation or the smallest window, where, as I mentioned it here as point (1)–the library and the work–of course you can say that if you're able to exchange your components then maybe your library and your work will work, but maybe not the entire hardware, and I mentioned examples here under 4… So, maybe not the entire hardware unit, the computer or its peripherals, or the car, but the software, as is, or as you or the user looks at, he, that might work.

So, that's also a way of interpretation, and I would like to open the discussions on that, but you can also have that a little bit later–looking at it to understand, okay, how would you interpret the license in that way?

So, this brings me already to the end. I think, and I hope, that I showed you that security and safety is really of high importance for the car industry. We are in a very highly regulated market where the governments and the bodies look to the car because it's not a thing… The car is not something you just put in your pocket or you can use in your private environment at home, but a car is something which drives on the road and can be of a high danger to your body and your life, so it's heavily regulated.

Whenever, and I'm also very open on whatever I've been, as

Eben mentioned already, I've been following the whole discussion since 2004, and I always tell also the people in my organization that we cannot hold up all of that. We have to be on the track. I won't be part of the whole system because otherwise we might lose, as Eben said, we might lose the track to innovation. So, I don't want to hold back the whole thing, we have to be part of it.

I also agree that GPLv3, we shouldn't hold that up, we should look carefully how we can do that, how we can make it operate properly in a car, but I think–I don't want to prevent technological innovation out of our cars, and because I can also take, as I said the very safe way and say, "Okay, GPLv3 is absolutely not permitted in our cars," but just as a general rule I don't want to do that. Right? Because, at the end of the day, I have to go to our board and say, "Okay, this innovation, we are not able to bring that innovation to our cars because there's just a general rule which says, okay, it's not allowed."

So, I think we have to take a very differentiated approach to the topic and look at it case-by-case in order to be on the very top part of the innovation–to look at it and not to prevent innovation. And I think the more that people are able to look at a software the more that innovation will be there. Though there might be developers, very smart people, in our company, they might have a certain view on a technology or a software, but there might be even smarter people out there, and I think this is something we should use, and I'm trying to encourage my people in our organization to do that.

So, thank you very much for listening to me, and I hope for a fruitful discussion.

MOGLEN: Thank you, Daniel. What I think I would like to do is to get all the voices into the discussion, and then we can take all the questions from multiple angles.

I should just clarify before we move on that nobody ever sued TiVo about anything, what happened in this history was that TiVo made a digital video recorder for home use, which did, indeed, as you say lock down the entire software stack in the box and we made GPLv3 in the knowledge of the existence of that business model for the production of appliances, which it was the case that my client, Mr. Stallman, did not like.

So, we began making GPLv3 with a requirement that users be able to install modified versions of GPL software, and my client made anti-TiVo-ization the label under which that operated, which, of course, put a particular company in the headlights. I don't think that Donald Trump learned about tweeting at Amazon from Richard Stallman, indeed, I don't think Donald Trump has every learned anything from Richard Stallman, but to be a lawyer for a guy who is singling out particular companies raises certain difficulties.

I found myself one day in conversation with the general counsel of TiVo, Max Ochoa, he no longer works at Tivo, and Max said, "Look, you know, if you guys would agree to drop all this anti-TiVo-ization stuff, we would stop encrypting the movies on the hard drive." And I said, "Gee, Max. That won't help, we're not the free movie foundation. It's the Free Software Foundation. What we're concerned about is peoples' ability to modify the software in the device so they can fool around with making it work better for them." "Oh," he said, "We could never permit that because then there wouldn't take the program guide." I said, "You mean that Andrew Tridgell in Australia is going to modify his TiVo and he will decide to do without the program guide. That's one guy. Aunt Sally will never do it." "No," says Max, "You don't understand. We lose so much money on each piece of hardware we sell that if they don't take the program guide, even if one user doesn't take the program guide service from us, then we're out of business." I said, "Well, Max, look, here's the problem that we have: we make free software, and we do the very best we can, we give it to everybody to use for whatever reasons they want in any way they please, and we don't charge them. You're asking me to accept a terrible tax on our business model so that you can sell table-top super-computers below cost. This is not actually a really good outcome for either one of us. Selling hardware below cost is not a good long-term business, and putting us into deep trouble…"

So that's really in the end what TiVo-ization was about for TiVo. It was securing a service monopoly connected with hardware sold below cost–a twenty-first century business model that isn't very good and that BMW or Audi would never accept. Nobody will ever lose money on selling the car so that they can make a service for it. They want services, I grant you, but the car must be profitable.

So, my question, really, I think will turn out to be, what is it that is the stake in locking down all the software in the car, and can we help? What I think we have now seen is that what Mark and I are talking about is a version of your partitioning structure on steroids, meant to work one-thousand times better for you, and that we are really saying that we now have technology on the shelf for you that would allow you to achieve the kind of control that you want in a very highly potentiated way, so that you could both protect the things you need to protect and allow tinkering with the things that it wouldn't hurt you to allow, and, then, all of a sudden, the licenses would cease to mean very much to you because you would have the level of control that you would need over the technology in order to have the level of control you need in your business.

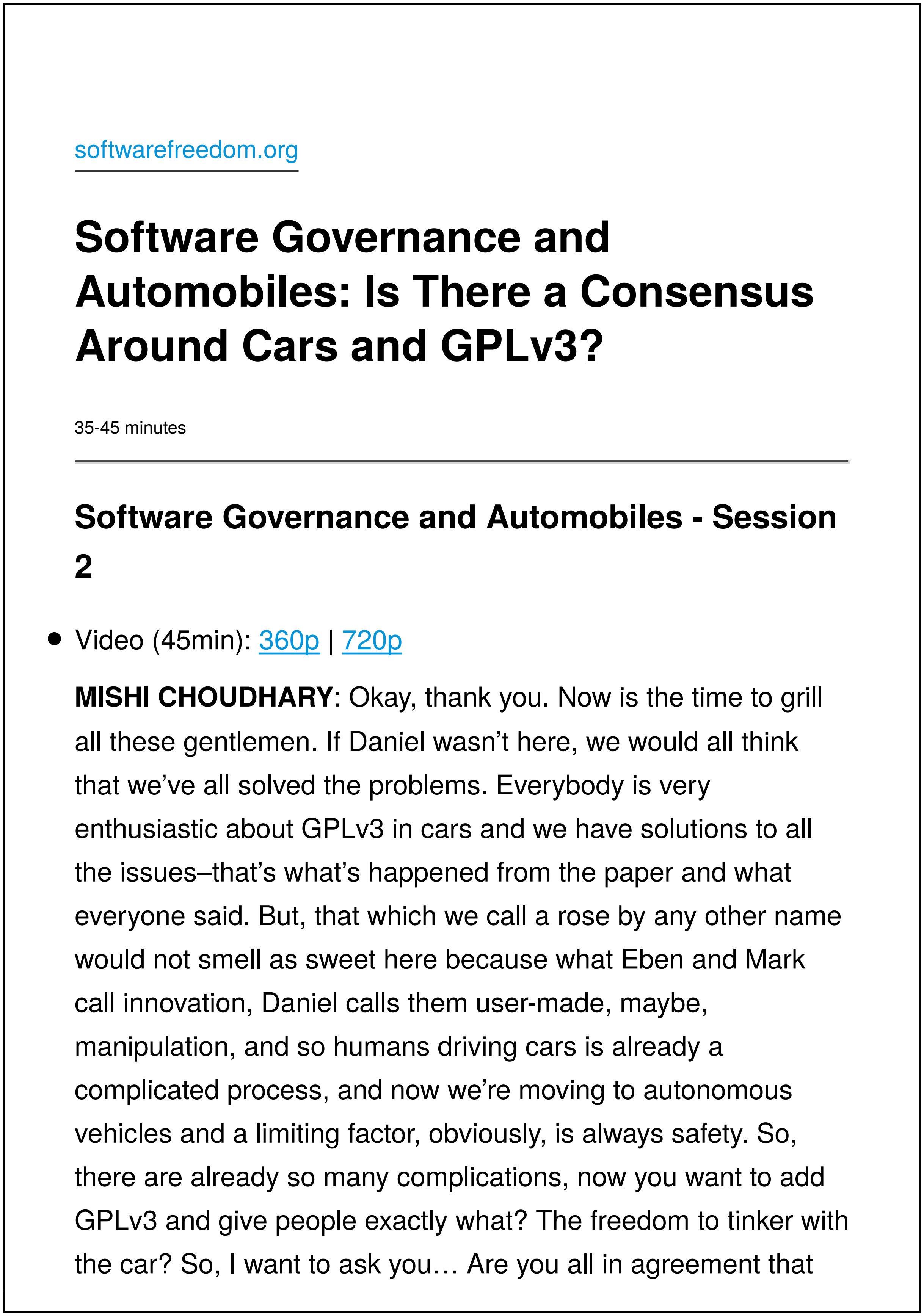
But the problem is TiVo-ization is an all-or-nothing idea–I lock it all down or I let it all out, and what we really need is very fine-grain stuff, and it's not in the language of the licenses, it's in the packaging of the software.

That's where I think the conversation is at this moment, which is why we really need Jeremiah because Jeremiah is the person who lives exactly in the middle of that discussion.

Previous: 1b-shuttleworth | Next: 1d-foster | Contents

Chapter 16

Is There Consensus Around Cars and GPLv3? (Choudhary, et al.)



softwarefreedom.org

Software Governance and Automobiles: Is There a Consensus Around Cars and GPLv3?

35-45 minutes

Software Governance and Automobiles - Session 2

- Video (45min): 360p | 720p

MISHI CHOUDHARY: Okay, thank you. Now is the time to grill all these gentlemen. If Daniel wasn't here, we would all think that we've all solved the problems. Everybody is very enthusiastic about GPLv3 in cars and we have solutions to all the issues–that's what's happened from the paper and what everyone said. But, that which we call a rose by any other name would not smell as sweet here because what Eben and Mark call innovation, Daniel calls them user-made, maybe, manipulation, and so humans driving cars is already a complicated process, and now we're moving to autonomous vehicles and a limiting factor, obviously, is always safety. So, there are already so many complications, now you want to add GPLv3 and give people exactly what? The freedom to tinker with the car? So, I want to ask you… Are you all in agreement that there is no future of cars without free and open source software? And I want you to talk about that agreement which obviously has a lot of disagreement built in. Daniel?

DANIEL PATNAIK: Yeah, that's what I wanted to point out during my presentation. I think there is a future of open source software in cars. This is a fact which I can see everyday. So, there is already open source software in the cars, and there is definitely a future. I remember and I just mentioned that–when we were standing together–some years ago, some people said, "Okay, we want to block that entirely," and I said, "Hang on a minute, we cannot and we should not do that, and this is not the way forward," so I was always encouraging people to take a very precise look at what we are talking about so we can enable, we can show the boundaries and enable software innovation to get to the cars within the boundaries that are important to us.

CHOUDHARY: Mark, what is trusting software? It's not just knowing the provenance of software, but it's also about what you talk about in the paper–about how software governance is managed. Daniel also talked about partitioning, so can you talk a little bit more about, in that context, the future of FOSS in cars, and how you see it?

MARK SHUTTLEWORTH: I guess I'm reminded of that old 1980s Cold War, "trust but verify," and I think, with hindsight, that was a pretty savvy view, that ultimately trust isn't a simplistic thing–it's best if it's backed by science, it's best if it's backed by facts, and possibly if it's backed by teeth as well. So, why do we trust something? Because we believe it will be a predictable outcome–we believe we can predict the outcome, and I think what I observe in the industry is that we're going through that gradient of going from trust is the sort of nebulous thing at a very high level that's almost tribal and branded, you know, and now we're getting down to a sharper, pointier more, almost more, useful definition of what do I need to trust and do I or do I not trust that for that, effectively.

CHOUDHARY: Daniel, you want to jump in and also talk about how you think this trust plays out, how the car really works, what you lock-down and what you keep open?

PATNAIK: Yeah, I think also here we have to look where we need trust. I think, if you look at the overall car, of course not only we want to have trust that everything works well but also the customer, of course, wants to have trust that everything works well. At the end of the day, also, the trust is part of the permission of the car and I think that this is one of the key issues here that are regulatory side as well. So, as long as they say there's trust, and also from a regulatory side, there is a permission to it. However, I think we have to distinguish, and I don't have a clear answer on that right now, but we have to distinguish where we really need trust and maybe we have more of a freedom to back up a little bit and say maybe here, in that area of the car, trust is not so necessary but in other areas trust is very, very important.

EBEN MOGLEN: Yes, the difficulty here is that trust is a different concept when you're not bending metal, you're making software. The way that the vehicle OEMs got trust in the

> physical automobile was by saying, “Please use only General Motors replacement parts, please use only our approved spark-plugs… Please this, please that,” and the idea was that somebody manufactures the trusted thing and then there’s a whole bunch of people out there who manufacture untrusted things. Please use the trusted objects and then your car will perform correctly, that’s not a twenty-first century concept anymore. Now, we have the problem of no software is ever perfect, therefore the idea that what you do is you manufacture a TiVo-ized car, you put some software in it, you lock the software down, and that software works until the car dies is never going to be correct. It’s never going to work that way. The problem with the idea of TiVo-ization is that it establishes trust at the moment the car leaves the factory, and now you are trusting all the defects in the software for the life of the car and nobody really believes that. Therefore, we are talking about an environment in which we’re going to have to have software-replaceable parts, and we’re having a discussion about who we trust to make and replace those parts. This is, in the end, trusting people not trusting software–trusting software is just a reflection of trust in persons.
>
> What the re-organization of Volkswagen reminds us of this week, yet again, is that the idea that the only trusted persons are the manufacturers is also not going to be correct in the twenty-first century. We talk about this highly regulated market, but we now understand that the Volkswagen case was extremely useful in this, too–that regulators are not going to find the problems in the software, civil society is going to find the

problems in the software. This is 100% guaranteed to be true; mathematically, it's true. The regulators will never employ enough people. There are not enough taxes in the world to employ enough regulators to check all the software–civil society is, therefore, going to be responsible for inspecting and discovering failures in software. That means FOSS by design because otherwise we're using unsafe, uninspectable building materials. And now the problem is so you have inspected and you have found a problem, then what? You write a letter to the automobile manufacturer and you say, "You guys ought to get around to fixing this one of these days?" You write to the National Highway Traffic Safety Administration and say, "I found a bug, would you please recall all these automobiles for me?" None of the existing mechanisms will work with respect to what is going to be the most complex, the most dangerous, and the most widespread bunch of software in civil society. We're going to have to figure out ways to govern repair, modify, and use that don't depend on trust in a brand on the side of spark-plug box.

CHOUDHARY: Daniel says take the principles, what FOSS teaches you, but not necessarily the license itself because openness comes in and all regulators would like some throat to choke when there is a problem, so…

MOGLEN: Yes, and it won't be a legal throat to choke, it isn't a copyright lawsuit against somebody. It's a technical set of facilities that operate software in vehicles in such a way that regulators, users, manufacturers, parts manufacturers, and third-party service entities can, together, optimize the mix of software in the vehicle at any given moment, given where it is and what it's doing, and that's going to turn out to require more sensitive mechanisms than either "free-flier" zone, it's all open, sometimes it gets fixed at annual check-up or fifty-thousand miles or a hundred thousand miles, or every Johnny and Sally makes whatever changes she wants to her automobile–neither one of those are going to be acceptable.

Somewhere in between there has to be a way of doing that more sensitively, and that has to be not a legal set of rules, it has to be a technical set of rules, supported by law where we can use contracts and copyright law and other legal machinery to keep everybody to it, but without acceptable technical solutions for the very complicated problem of governance, as a technical matter, we're not going to wind up with what we want.

If automobiles are TiVo-ized in the twenty-first century and nobody can change the software in them but the manufacturers, the manufacturers are going to wind up very unhappy… Because they're going to be responsible for a nightmare of liability problems as software ages and conditions change and they're the only people who can fix it. One of the reasons that I think it's so important to talk about these liability issues is that I think at the moment the automotive manufacturers think that the best way to avoid liability is to control it all themselves, thus piling up all the liability in their hands, and I think what Mark and Jeremiah and I are all saying in different ways is that's not the right long-term solution. It doesn't optimize innovation. It doesn't optimize liability protection. It requires you to be vertically integrated servicers of long, lifetime safety critical software forever. Are you really sure you want to be in that business? So, my question is, does your client really want to be in that business? Does your client, Audi, really want to be in that business of centralizing in itself all liability for software problems forever and being the only point of repair for TiVo-ized software in cars?

PATNAIK: Good question. I think, of course, to a certain extent we want to and have to control a big bunch of it as long as we don't have real solutions because liability is probably two-fold. The liability, as you mentioned, I can understand, and I see that point as well, but also, from a product liability side, the government or the courts they have also some obligations on the manufacturers–you have to look at what is been done with your car, with your car system, with your software, and if you see that someone is doing something you will also have to control this or, at least, do analysis of that and you are also responsible in a certain way to ensure that whatever is being done with your software is safe. So, this is the context we are in. So, of course, the question is do we all have to be–do we all want to be–liable for everything? No, probably not. If there's a way to divest that, of course we will be open to do that, but, I think, we are not yet there. We have to get there.

MOGLEN: So there are two things we could think about with respect to that. The first thing is that the most dangerous thing that a human being can do to modify their car is to make uneven the inflation pressures of their tires, and nobody would ever say that General Motors is going to be responsible because Jimmy decided that he liked fifteen pounds per square inch in the left rear tire and forty-five pounds per square inch on the front passenger tire. The resulting wrapped-around-a-telephone-pole experience is not regarded as the manufacturer's problem. There was a user modification, and it was deadly. It's way more dangerous than screwing with the VLC that plays the entertainment video in the backseat for your kids so that the volume control will never go past four, which I predict will be a popular modification, right?

We need to understand the scale at which what we're trying to do is figure out what forms of software modification are actually not a problem after we have given the manufacturer tools that allow it to control the stuff that really is a problem quite heavily while not controlling that which is not, and no license can do that, no bunch of legal words can do that, there has to be a technical infrastructure, connected with the way software is put together and distributed, that gives us that.

What Mark is saying about Ubuntu core as another addition of the software is that the people who have been most innovative in distributing FOSS in the last generation, and who've changed the software industry around the world by doing it, are now concentrated on that question… Because of IoT, because of the automobile, because of all that complex stuff at the edge, we're now going to learn to package and distribute software with all of that kind of sensitive control…

I want to find a way of bridging the remaining legal difficulties, whether it's GPLv3 or your concerns about how LGPLv2.1 works or–I want to take all of that legal material and re-shape it just a little bit around the edges so that we can understand how it works compatibly with new packaging structures. To give manufacturers fine-grained enough control that they can relax their concern about user modification.

They live with the fact that you can't control tire inflation pressures from the moment the car leaves the factory. They know that there's no TiVo-izing the pressure valves in the tires, they understand that there's no way that every single thing can be controlled in the interest of safety and in the interest of liability limitation. But there are obvious things that we would like to be able to do, including to have a computer in the car which constantly monitors tire pressures and that puts a note up on the dashboard if something is wrong, right? Which is software that we might allow people to modify but we might also allow them to modify it only in certain ways, and we would certainly want to have control over provenance. You don't want me modifying your tire pressure gauges in your car with an over-the-air modification, and this, again, is one of the things that Mark and I are trying to address in the paper–to explain how we can use digital signatures and blockchain publication and other things so that everybody, NHTSA after a crash and third-party manufacturers and police investigators, can all know exactly what was the software in the car and who put it there.

That's going to be a critical part of trust–and a critical part of law–in the twenty-first century. If somebody made a modification to a Tesla's auto-pilot software and it wound up in the middle of a median divider on a highway at seventy-five miles an hour, who changed those bits is going to be a very important story.

PATNAIK: I fully agree, Eben. I fully agree with your point of view. We have to find ways to be able to have a differentiated view to certain things, and we have to think about how we can get there. I think I put up some ideas about how we could get there. I think you, Mark and Eben, have also showed a way. I think we have to be open, and I would like to–I think it's a good discussion to be open-minded to understand how we can get there in order to accomplish all of that.

MOGLEN: Maybe we should see who else has questions…
Mr. McGuire?

CHOUDHARY: Sure. Nicholas.

NICHOLAS MCGUIRE: Before you go into user-modified–or the problem of user-modified–what is the expected modification rate of the OEMs? And that is so extraordinarily high already that you, with your current model, can't even handle that, and that's why I think the discussion about is it the traditional model versus the user-modified model is actually the wrong discussion. The discussion that we need for the automotive industry is the traditional, "I control the software," versus "I have a highly dynamic software that I'm going to be updating probably something like every two weeks, once I have the complexity of an autonomous vehicle," and if you can solve that problem, extending that for user-modifiability will be significantly easier. But as long as we discuss from these two very far apart sides, I don't think we're going to get close.

PATNAIK: It's difficult to answer because I didn't even clearly get the question, but of course we have these two-sided views… How I understood yourself was that we should combine the two things, but… Yes, I think, of course, there will be updates and probably even regular, if you see it with an everyday device–everyone is using today, there will be updates every second day, every day, and this will happen, also, with cars in the future, the more software will get incorporated in our cars… Still, however, I think we should–and we have to differentiate a little bit about what the individual user does, coming back to Eben's example: if the customer modifies his individual car and his tire pressure, of course he can do so and it's his own risk and I don't want to prevent him, though I would like to in some cases, but this is his personal decision to do so, but if he then makes this public and gives this solution to other people, then I think we are in a position and we at least have to not control it but at least know that and have a position for how we react to that.

SHUTTLEWORTH: There's a reference earlier, sort of tangentially to regulation and liability, and what's interesting for me is that this narrative often plays out as a contest of wills between the private individual and the institution and the institution's commercial interests, but the really interesting cases are all, typically, regulated. And so, the balance of interests is much more complex than just a private individual and an institution, and, I guess, in a sense all of our interests are represented in the regulatory function–all of our interests are represented in civil behaviors and decisions.

I think it's really important that we figure out mechanisms to represent those stories. We may well come to the view that actually that is very helpful to manufacturers because it essentially starts to establish the limits of their control or the limits of their expectation of control and, therefore, liability.

Anything that essentially is on public roads or anything that is essentially in a public environment becomes something like a shared responsibility and having a clear limitation on what you're expected to enforce, potentially, is helpful in the bigger picture.

MOGLEN: Mr. McGuire's point that the amount of modification occurring in the software in vehicles is going to be extremely high, all the time, I think is unanswerably correct, right? Once we are talking about software doing the driving, it is going to be updated all the time, to take account for all kinds of experience that was unexpected, and even before we get to that…

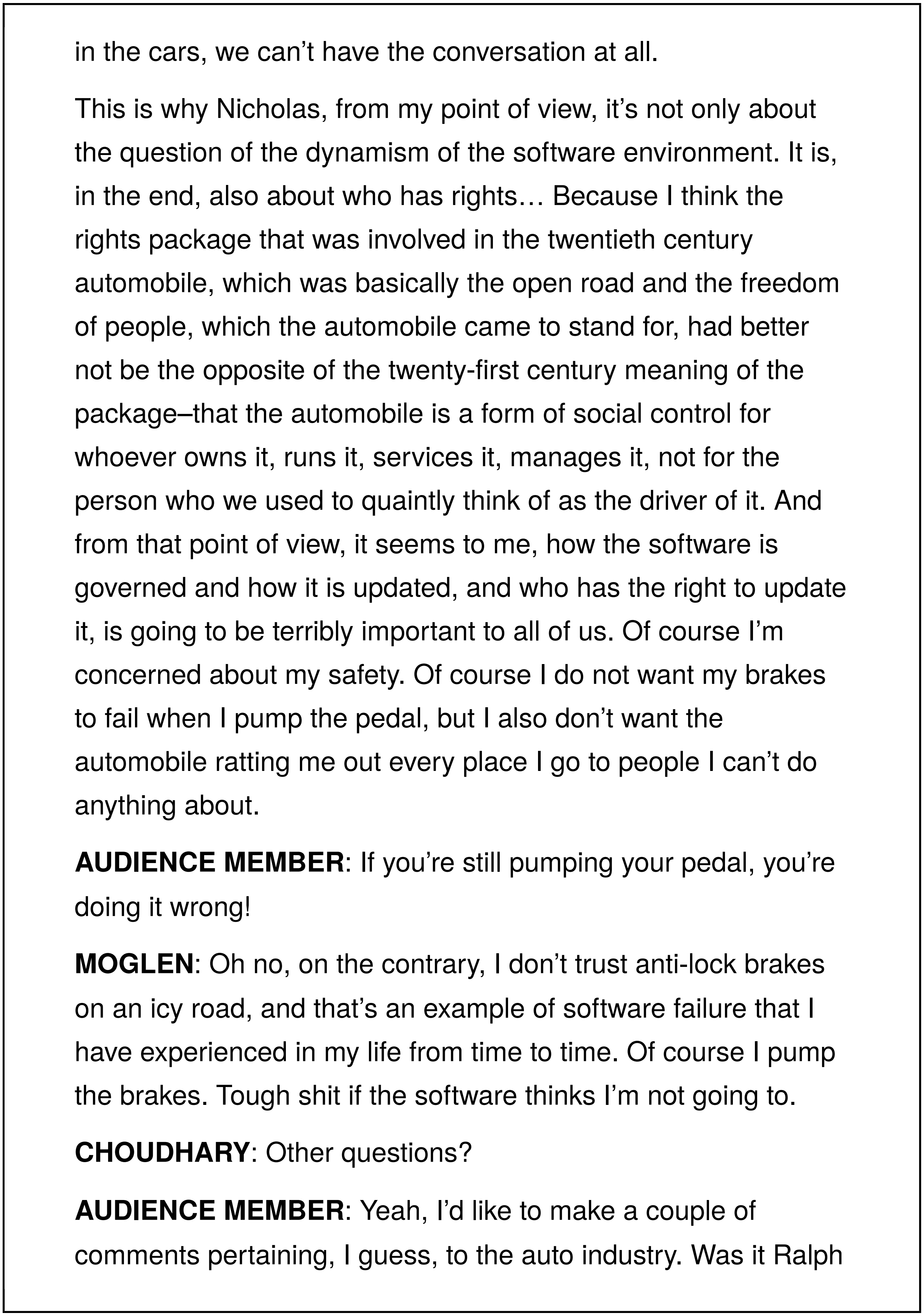
Look, there is an argument that the anti-lock breaking software, which we all want to think of as, you do it once, and you do it right, and you never let anybody change it again, really ought to be changed according to weather conditions and all sorts of other subjects and that we really ought to want a high degree of software volatility inside the vehicle, but without strong governance principles, including the ability to roll back halfway, we're going to wind up in a world where automobiles that can kill people are no more successful completing their updates than Windows 10 objects, which, after all–pardon me, Justin, it's hard for me to imagine that I have friends in Microsoft, I need to be careful about what I say about them in public, but let's face it that even in a comparatively simple environment called, "one device we provide the operating system for," high dynamism in updating is a terrible, terrible, problem and we need better tools for it.

I also think that something Jeremiah said which is critically important at this moment, when we talk about regulation. In the world that the governments are now looking at it is the data generated around the car which is the greatest and most important subject. All this other stuff is comparatively traditional.

Now, there are two ways of thinking about that: one of which is that all the data generated around the vehicle is going to regulated and government controlled and the other way is my way, which is that better not happen. And one of the most important elements in what users of automobiles and other vehicles and autonomous systems in the twenty-first century are going to want the power to modify is the leakage of the data. I'm okay with my car having as much tendency to be chatty about who I am and where I'm going as is minimally necessary in order to achieve certain agreed upon social goals, and after that…? Right?

I mean, I live a reasonably effective life in the net without a Google account, without a Twitter account, without a Facebook account, without a platform relationship of any kind. Please don't tell me that in order to own an automobile in the twenty-first century I'm going to have to be more risky with my personal data and the substance of my life than I am already.

And that surely means that there are going to be levels of desire for user control over the way software in vehicles work which are extremely valuable to the individual, extremely important to manufacturers and service platforms, and extremely interesting to government regulators. The rules about all of that have to be adaptable. We have to be able to have that social policy conversation in a serious way, and without some kind of technology for governance of user-modifications of the software



in the cars, we can't have the conversation at all.

This is why Nicholas, from my point of view, it's not only about the question of the dynamism of the software environment. It is, in the end, also about who has rights… Because I think the rights package that was involved in the twentieth century automobile, which was basically the open road and the freedom of people, which the automobile came to stand for, had better not be the opposite of the twenty-first century meaning of the package–that the automobile is a form of social control for whoever owns it, runs it, services it, manages it, not for the person who we used to quaintly think of as the driver of it. And from that point of view, it seems to me, how the software is governed and how it is updated, and who has the right to update it, is going to be terribly important to all of us. Of course I'm concerned about my safety. Of course I do not want my brakes to fail when I pump the pedal, but I also don't want the automobile ratting me out every place I go to people I can't do anything about.

AUDIENCE MEMBER: If you're still pumping your pedal, you're doing it wrong!

MOGLEN: Oh no, on the contrary, I don't trust anti-lock brakes on an icy road, and that's an example of software failure that I have experienced in my life from time to time. Of course I pump the brakes. Tough shit if the software thinks I'm not going to.

CHOUDHARY: Other questions?

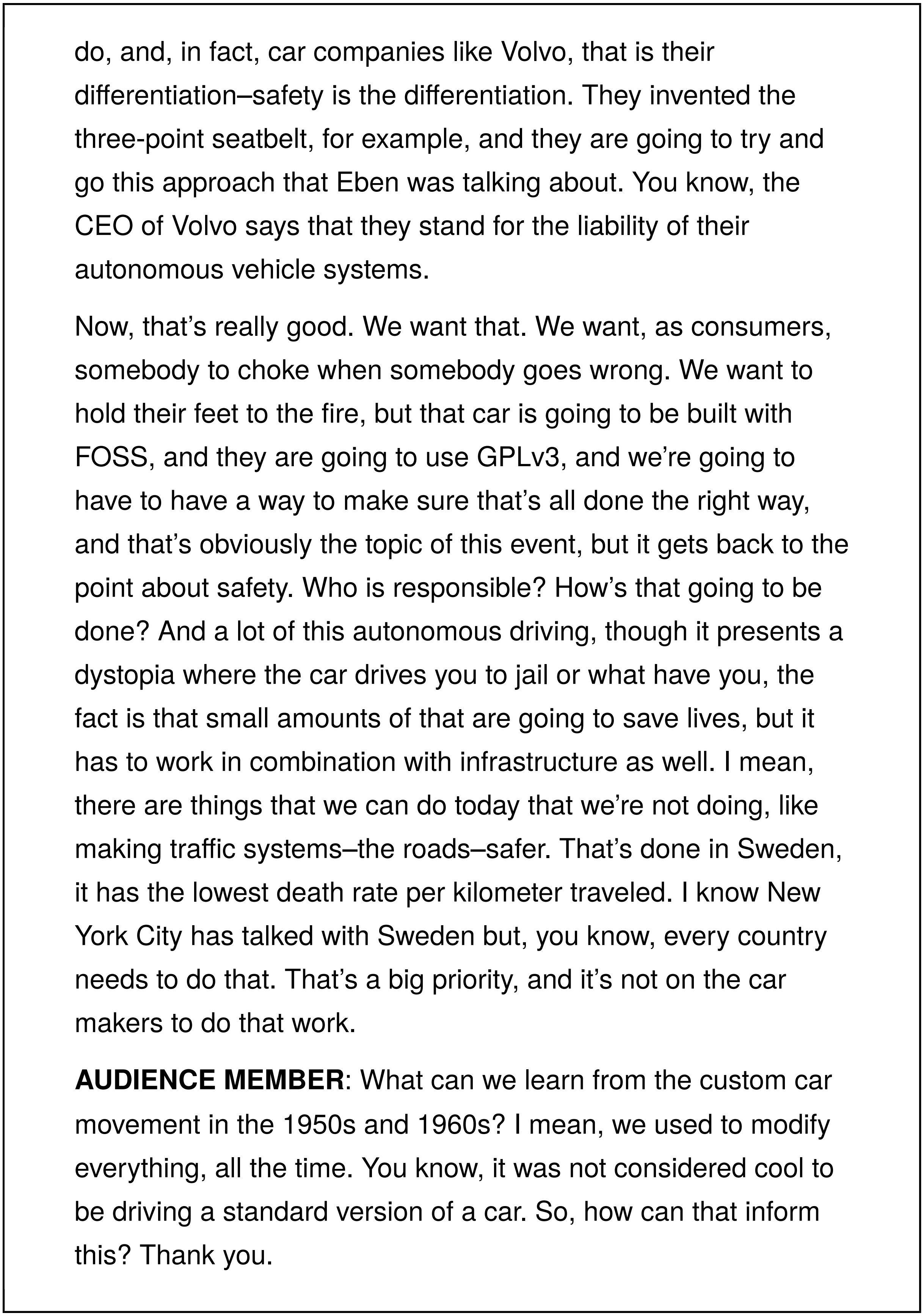
AUDIENCE MEMBER: Yeah, I'd like to make a couple of comments pertaining, I guess, to the auto industry. Was it Ralph

Nader who wrote the book, “Unsafe At Any Speed,” about the Corvair and the garbage that General Motors produced many years ago and, apparently, continues to produce today. I’m wondering about when the cell phones came along and we see a spike in auto accidents and followed by auto deaths, and I believe that last year auto deaths in the U.S. were approximately thirty-four thousand, so I wondered with the automotive manufacturers loading up distraction device after distraction device on the dashboard where you can watch a video, tune into the internet, and, generally, get distracted… So the auto deaths… You know, do the manufacturers really give a damn? And I’d say not really. And, then, in terms of the auto industry moving to Mexico, South Korea, and China, and then when you look at the J.D. Power’s quality control study of autos, I think there’s one U.S. automaker in the top ten–fortunately, Audi is in the top ten along with Lexus, Toyota, Honda, Nissan, etc… So, I guess my last point is to Daniel: Can you comment about the fact that BMW and Mercedes have recently announced a joint venture–I believe it’s really to counter the power of Google, Tesla, and Apple, where the automobile becomes the software machine on wheels, and the Germans don’t want to be squeezed out of the business by Silicon Valley. So, over to you, Daniel.

PATNAIK: Yes, thank you. It was a bunch of remarks and questions, but maybe to start with the last, I’ve been talking to the big players in Silicon Valley as well because I’ve been legally involved as well, and I know from our departments that we are, of course, doing the same. We are trying to… We understand that there are interests to come up with software solutions, we're trying to match and match those interests with coming up with our own solutions–it's either in the automotive industry with other automotive companies or within the group where we have more than thirteen brands, at least thirteen brands, at the moment. We have a big power and an interesting power in order to be competitive to the others, so we don't have to fear it, we have to watch that very closely.

But you mentioned also in the very beginning a part which said that distraction in the car is of–you mentioned that the producers don't give a damn, and, on the contrary, we give a lot of it. We know, in my department, at least, we have engineers that care a lot about product liability, and I work with them very closely, and distraction in the car is of a high importance, at least in our company. So, we don't allow things to–movie, as you said–we don't allow that. So, whenever the car starts moving, everything is shut down. Of course, I cannot prevent a customer to put his phone somewhere in his field, to attach it somewhere and to watch that–I cannot prevent that, but we are even looking at it and we're trying to see if there is a cable connected to try and stop that. But this is a very wide field that we can discuss about for a long time, but we care a lot about those issues–the customer to not get distracted by whatever is there.

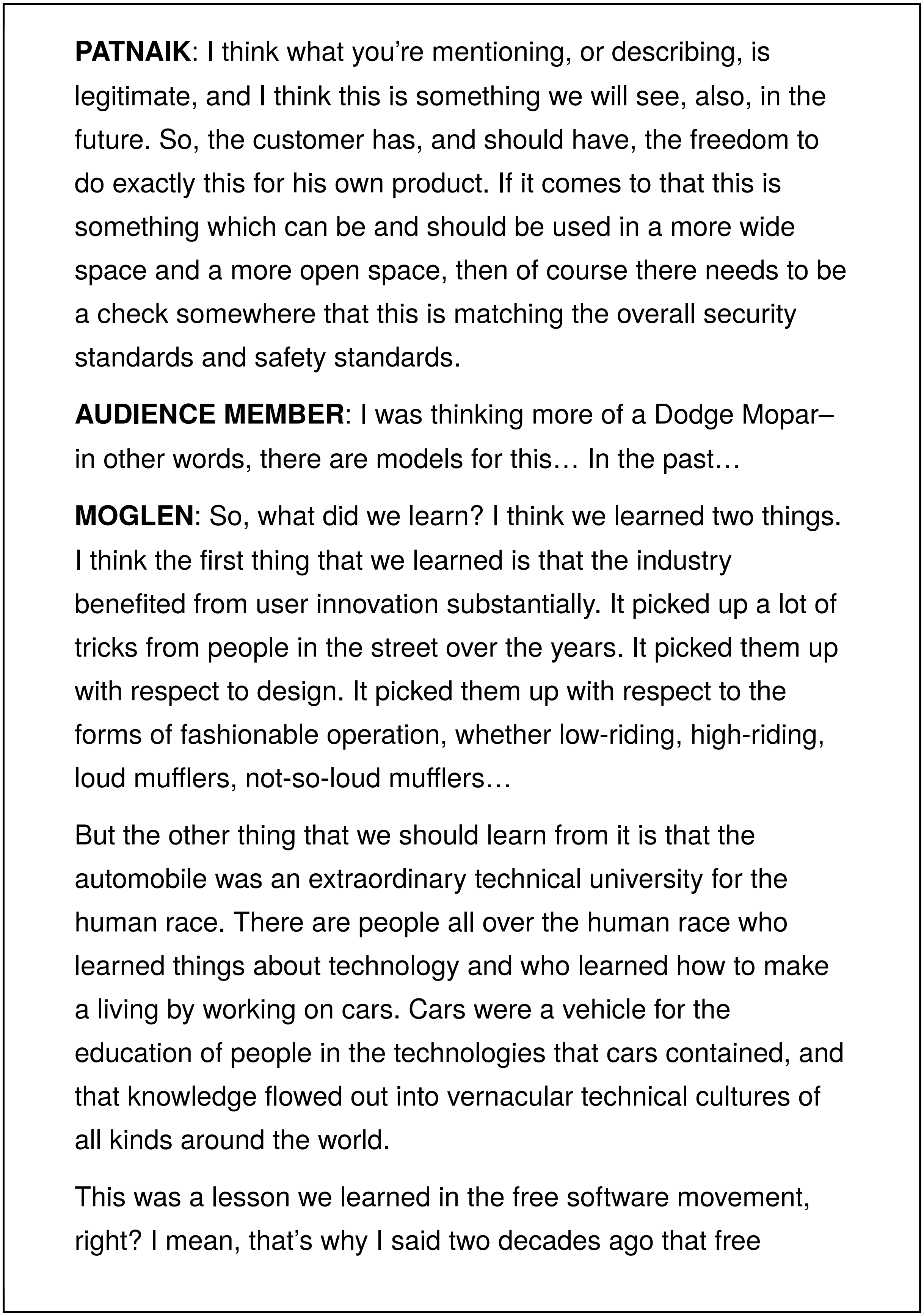
JEREMIAH FOSTER: Many car companies, I think, all care about safety. I think they care a lot more than we imagine because I think the message that gets out to society is much more commercial, it's much more selling, and I think that they realize they need to adjust their message. But they absolutely



do, and, in fact, car companies like Volvo, that is their differentiation–safety is the differentiation. They invented the three-point seatbelt, for example, and they are going to try and go this approach that Eben was talking about. You know, the CEO of Volvo says that they stand for the liability of their autonomous vehicle systems.

Now, that's really good. We want that. We want, as consumers, somebody to choke when somebody goes wrong. We want to hold their feet to the fire, but that car is going to be built with FOSS, and they are going to use GPLv3, and we're going to have to have a way to make sure that's all done the right way, and that's obviously the topic of this event, but it gets back to the point about safety. Who is responsible? How's that going to be done? And a lot of this autonomous driving, though it presents a dystopia where the car drives you to jail or what have you, the fact is that small amounts of that are going to save lives, but it has to work in combination with infrastructure as well. I mean, there are things that we can do today that we're not doing, like making traffic systems–the roads–safer. That's done in Sweden, it has the lowest death rate per kilometer traveled. I know New York City has talked with Sweden but, you know, every country needs to do that. That's a big priority, and it's not on the car makers to do that work.

AUDIENCE MEMBER: What can we learn from the custom car movement in the 1950s and 1960s? I mean, we used to modify everything, all the time. You know, it was not considered cool to be driving a standard version of a car. So, how can that inform this? Thank you.



PATNAIK: I think what you're mentioning, or describing, is legitimate, and I think this is something we will see, also, in the future. So, the customer has, and should have, the freedom to do exactly this for his own product. If it comes to that this is something which can be and should be used in a more wide space and a more open space, then of course there needs to be a check somewhere that this is matching the overall security standards and safety standards.

AUDIENCE MEMBER: I was thinking more of a Dodge Mopar– in other words, there are models for this… In the past…

MOGLEN: So, what did we learn? I think we learned two things. I think the first thing that we learned is that the industry benefited from user innovation substantially. It picked up a lot of tricks from people in the street over the years. It picked them up with respect to design. It picked them up with respect to the forms of fashionable operation, whether low-riding, high-riding, loud mufflers, not-so-loud mufflers…

But the other thing that we should learn from it is that the automobile was an extraordinary technical university for the human race. There are people all over the human race who learned things about technology and who learned how to make a living by working on cars. Cars were a vehicle for the education of people in the technologies that cars contained, and that knowledge flowed out into vernacular technical cultures of all kinds around the world.

This was a lesson we learned in the free software movement, right? I mean, that's why I said two decades ago that free software is the greatest technical reference library ever assembled on Earth because it's the only way that a person without skill yet in the art can learn from the very best that is being done in all the ways that are being done just by reading stuff you can get for free.

We want the automobile to continue to be the seed of vernacular technical education in the world. We can't do that if people can't modify the car.

This was the point about GPLv3's anti-lockdown provision in the first place. We were trying to preserve what we understood to be the way people in the world became technically highly capable–namely, by hacking on their own things, and we did not want the level of things in the world locked down that young people couldn't learn to program on to go up too high. That always seemed to us a global north-south issue. In the north, there were lots of people, if they bought one computer that was locked down they could by another computer that was not locked down, Linus Torvalds was a good example–it didn't bother him, he didn't need to worry about it. He could buy another computer. But all over the world, there are people who have exactly one, and it would be good if they could hack on it because that's where they're going to learn.

This should also be true about the car in the twenty-first century because we saw how important it was in the twentieth. So, for me, the stakes in the modifiability of the technology include all the human learning that will flow from that, which is a vast welfare lost to the world if the machines are not willing to allow people to learn from them, and that means ability to read and to understand but it also means the ability to experiment.

Sure, it's more dangerous when it's an object that travels at high speed, and, therefore, it would be really good if we were clever about it like the GPS in the car tells the software agent, "No more modifications now, he's on a smart-road." He has to be totally in-sync with all that complex built environment around him–or, "nah, he's in the middle of the back of beyond on his own real estate in a rural county, let him do whatever the hell he wants to do."

And to have the ability to move back and forth between a highly regulated software state on a smart road and a less regulated software state somewhere else–all of that lies within our existing technical capacity. So, if we have the technical capacity to do those things, we should… Because what I think we learned from the twentieth century history of the car was that technical enablement was really valuable to people–it made an enormous difference in peoples' intellectual and economic development and in their lives.

As I've traveled around the world in the last lifetime of mine, I have certainly seen an awful lot of things that were done by modifying cars that car industries learned from and, more importantly, that people learned from, and whether it was on a Caribbean island where everybody was using left-drive cars for right-side-of-the-road driving or whether it was the adaptations of the self-worked-out propane conversions in countries in the global south or whether it is the miracle of the auto-rickshaw–we may not want to ride in them but it's a miracle that they exist and they stay on the road decade after decade with guys doing fixes at the end of the day and putting stuff together with spit and baling-wire–all of that comes from the history you're talking about, that the automobile was a highly sophisticated, very complex, but also very enabling technology that people interacted with in a whole bunch of ways that we now call hacking, and it worked really, really well.

SHUTTLEWORTH: I'm sorry, you described a brand–was it like a Dodge…?

AUDIENCE MEMBER: Mopar.

SHUTTLEWORTH: Mopar? And was that sort-of a modified…?

AUDIENCE MEMBER: It was called an after-market… It was an entire after-market industry…

SHUTTLEWORTH: Right. So I have to ask because I'm a lot younger than I look, and my memories of the 1950s and 1960s are entirely manufactured from watching movies made before I was born, but my impression is that this was the first time when pretty much every family got access to a car and cars were super cool–it was still just a little exclusive but not really that exclusive, and what you're describing reminds me so much of the importance of tapping into passion, tapping into enthusiasms, and this is true for every brand. It's easy to forget once you've become successful, you think that it's you that makes you successful, but it's not really, right? It's peoples' passion for what you mean to them and so on.

So, we see this, in our little way, in the existence of derivatives of Ubuntu, right? People who have different passions to us but it's easiest for them to express those passions starting with

Ubuntu, and we just grant them the rights to do that because it costs us nothing and the reality is it's interesting what they do–it's much more interesting what they do, often, than anything I might do in a day, and it generates enthusiasm, it generates activity.

I can see real value for that in, for example, a car or another object manufacturer being able to say, "Look, as long as I can bound these things, I don't mind allowing the creation of a Mopar," right? An enthusiast's sort of derivative, effectively. As long as I can bound the pieces where I may have a regulatory issue, effectively. That is actually a huge asset to me because all of that time, all of that energy, all of that thinking, is effectively much more directly applicable to me than it is to any of my competitors, and so, we may well see–you know, as soon as we have the ability to start drawing these distinctions–that that kind of fan club enthusiasm is both enabled and encouraged.

AUDIENCE MEMBER: Can I just ask which regulator has agency in all of this and what's being done to ensure that there's more consistency across regulations in the various sectors?

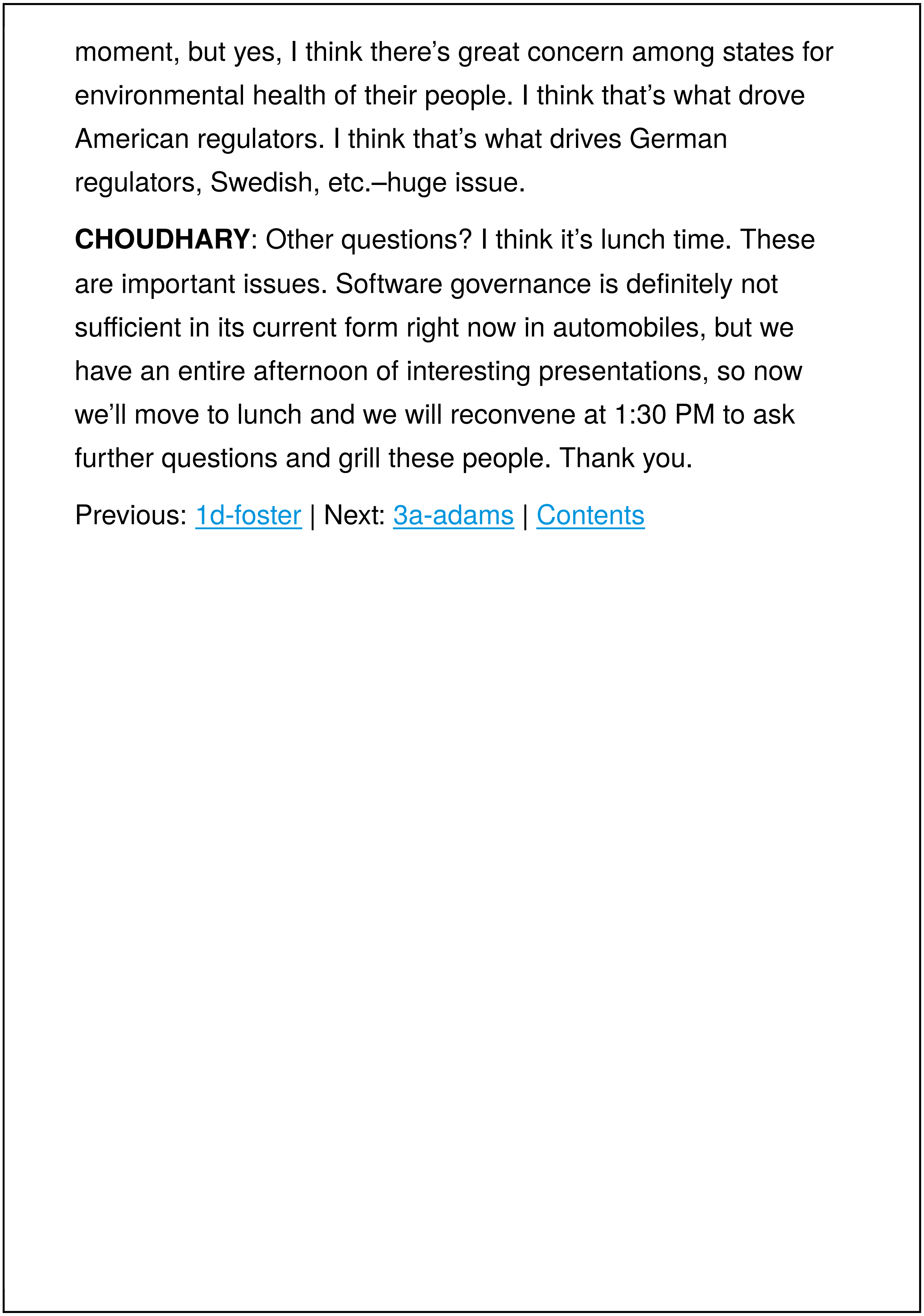
FOSTER: All of them. Yes and no. I don't think that there's a single governing body. In fact, part of the issue is that you have, for example, right to repair laws in Massachusetts that don't necessarily exist in other states. You have CARB, the California Air Regulatory Board, which basically sets policy for the United States when the federal government is not fighting them doing so, and then you have jurisdictions across the rest of the world, which may match or may be completely opposed, and then you have governments that both want to create new regulations for new income streams as well as preserve a differentiation or an opportunity for their own automotive industries to be competitive, so, yeah, that kind of harmonization I don't see it existing any time.

PATNAIK: I just added, I don't see that as well. They're all multi-national. Every national authority has its own view to that, and I don't see that there's something being done in order to bring everything together, but I would like to see the eyes of the regulator in whatever nation we look now. If we tell him, "Okay, we have produced and certified a car, and we have allowed the customer to do whatever he wants with the car and he, by the way, he's just driving down the road here." I want to see the eyes of the regulator, so I think there's a lot to do and a lot to discuss on the side of the regulators to also to make them understand the issue of FOSS being used in cars.

CHOUDHARY: If there's a regulator in the audience, this is your time. Other questions?

AUDIENCE MEMBER: Yeah, quick follow up for Jeremiah. Jeremiah, you mentioned the safety of Volvo, well, as you well know, Volvo is now a Chinese company that has said that they're going to switch over to electronic vehicles completely–is it in 2022 or something like that? So, again, we have the Chinese to look to for a global technology leadership in a green environment with less CO2 put into the atmosphere. Is Mr. Trump listening? Less CO2 into the atmosphere because we don't want to have to move to Mars.

FOSTER: Yes, I think Mr. Trump has other concerns at the



moment, but yes, I think there's great concern among states for environmental health of their people. I think that's what drove American regulators. I think that's what drives German regulators, Swedish, etc.–huge issue.

CHOUDHARY: Other questions? I think it's lunch time. These are important issues. Software governance is definitely not sufficient in its current form right now in automobiles, but we have an entire afternoon of interesting presentations, so now we'll move to lunch and we will reconvene at 1:30 PM to ask further questions and grill these people. Thank you.

Previous: 1d-foster | Next: 3a-adams | Contents

Chapter 17

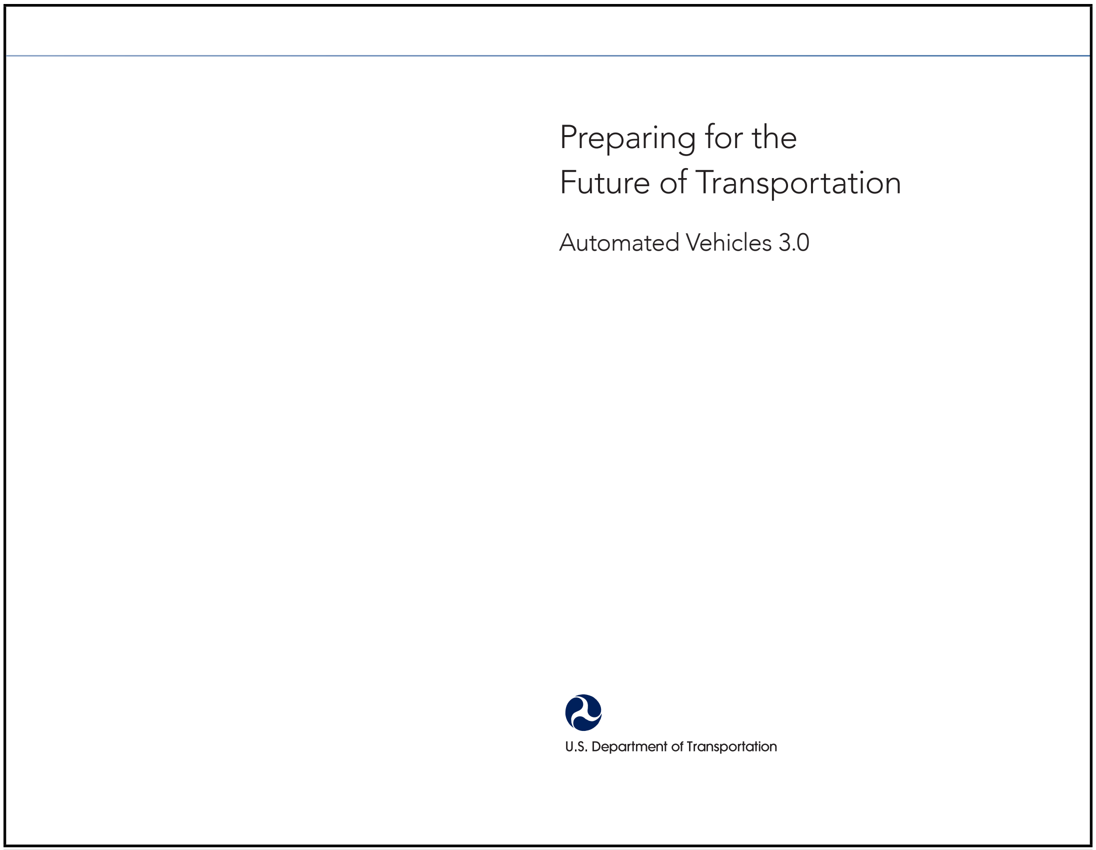
Preparing for the Future of Transportation (U.S. Dept. of Transportation)

U.S. Department of Transportation

Automated Vehicles 3.0

PREPARING FOR THE FUTURE OF TRANSPORTATION

With the development of automated vehicles,
American creativity and innovation hold the potential to
once again transform mobility.



Preparing for the Future of Transportation

Automated Vehicles 3.0

U.S. Department of Transportation

LETTER FROM THE SECRETARY

Secretary Elaine L. Chao
U.S. Department of Transportation

America has always been a leader in transportation innovation. From the mass production of automobiles to global positioning system navigation, American ingenuity has transformed how we travel and connect with one another. With the development of automated vehicles, American creativity and innovation hold the potential to once again transform mobility.

Automation has the potential to improve our quality of life and enhance the mobility and independence of millions of Americans, especially older Americans and people with disabilities.

Moreover, the integration of automation across our transportation system has the potential to increase productivity and facilitate freight movement. But most importantly, automation has the potential to impact safety significantly—by reducing crashes caused by human error, including crashes involving impaired or distracted drivers, and saving lives.

Along with potential benefits, however, automation brings new challenges that need to be addressed. The public has legitimate concerns about the safety, security, and privacy of automated technology. So I have challenged Silicon Valley and other innovators to step up and help address these concerns and help inform the public about the benefits of automation. In addition, incorporating these technologies into our transportation systems may impact industries, creating new kinds of jobs. This technology evolution may also require workers in transportation fields to gain new skills and take on new roles. As a society, we must help prepare workers for this transition.

The U.S. Department of Transportation is taking active steps to prepare for the future by engaging with new technologies to ensure safety without hampering innovation. With the release of *Automated Driving Systems 2.0: A Vision for Safety* in September 2017, the Department provided voluntary guidance to industry, as well as technical assistance and best practices to States, offering a path forward for the safe testing and integration of automated driving systems. The Department also bolstered its engagement with the automotive industry, technology companies, and other key transportation stakeholders and innovators to continue to develop a policy framework that facilitates the safe integration of this technology into our transportation systems.

Preparing for the Future of Transportation: Automated Vehicles 3.0 (AV 3.0) is another milestone in the Department's development of a flexible, responsible approach to a framework for multimodal automation. It introduces guiding principles and describes the Department's strategy to address existing barriers to safety innovation and progress. It also communicates the Department's agenda to the public and stakeholders on important policy issues, and identifies opportunities for cross-modal collaboration.

The Department is committed to engaging stakeholders to identify and solve policy issues. Since the publication of *Automated Driving Systems 2.0: A Vision for Safety*, the Department has sought input on automation issues from stakeholders and the general public through a wide range of forums including formal Requests for Information and Comments. In March 2018, I hosted the Automated Vehicle Summit to present the Department's six Automation Principles and discuss automation issues with public and private sector transportation stakeholders across every mode. The ideas and issues raised by stakeholders through these forums are reflected in this document. The goal of the Department is to keep pace with these rapidly evolving technologies so America remains a global leader in safe automation technology.

AV 3.0 is the beginning of a national discussion about the future of our surface transportation system. Your voice is essential to shaping this future.

E. L. Chao

Working together, we can help usher in a new era of transportation innovation and safety, and ensure that our country remains a global leader in automated technology.

U.S. DOT AUTOMATION PRINCIPLES

The United States Department of Transportation (U.S. DOT) has established a clear and consistent Federal approach to shaping policy for automated vehicles, based on the following six principles.

1. We will prioritize safety.

Automation offers the potential to improve safety for vehicle operators and occupants, pedestrians, bicyclists, motorcyclists, and other travelers sharing the road. However, these technologies may also introduce new safety risks. U.S. DOT will lead efforts to address potential safety risks and advance the life-saving potential of automation, which will strengthen public confidence in these emerging technologies.

2. We will remain technology neutral.

To respond to the dynamic and rapid development of automated vehicles, the Department will adopt flexible, technology-neutral policies that promote competition and innovation as a means to achieve safety, mobility, and economic goals. This approach will allow the public—not the Federal Government—to choose the most effective transportation and mobility solutions.

3. We will modernize regulations.

U.S. DOT will modernize or eliminate outdated regulations that unnecessarily impede the development of automated vehicles or that do not address critical safety needs. Whenever possible, the Department will support the development of voluntary, consensus-based technical standards and approaches that are flexible and adaptable over time. When regulation is needed, U.S. DOT will seek rules that are as nonprescriptive and performance-based as possible. *As a starting point and going forward, the Department will interpret and, consistent with all applicable notice and comment requirements, adapt the definitions of "driver" and "operator" to recognize that such terms do not refer exclusively to a human, but may in fact include an automated system.*

4. We will encourage a consistent regulatory and operational environment.

Conflicting State and local laws and regulations surrounding automated vehicles create confusion, introduce barriers, and present compliance challenges. U.S. DOT will promote regulatory consistency so that automated vehicles can operate seamlessly across the Nation. The Department will build consensus among State and local transportation agencies and industry stakeholders on technical standards and advance policies to support the integration of automated vehicles throughout the transportation system.

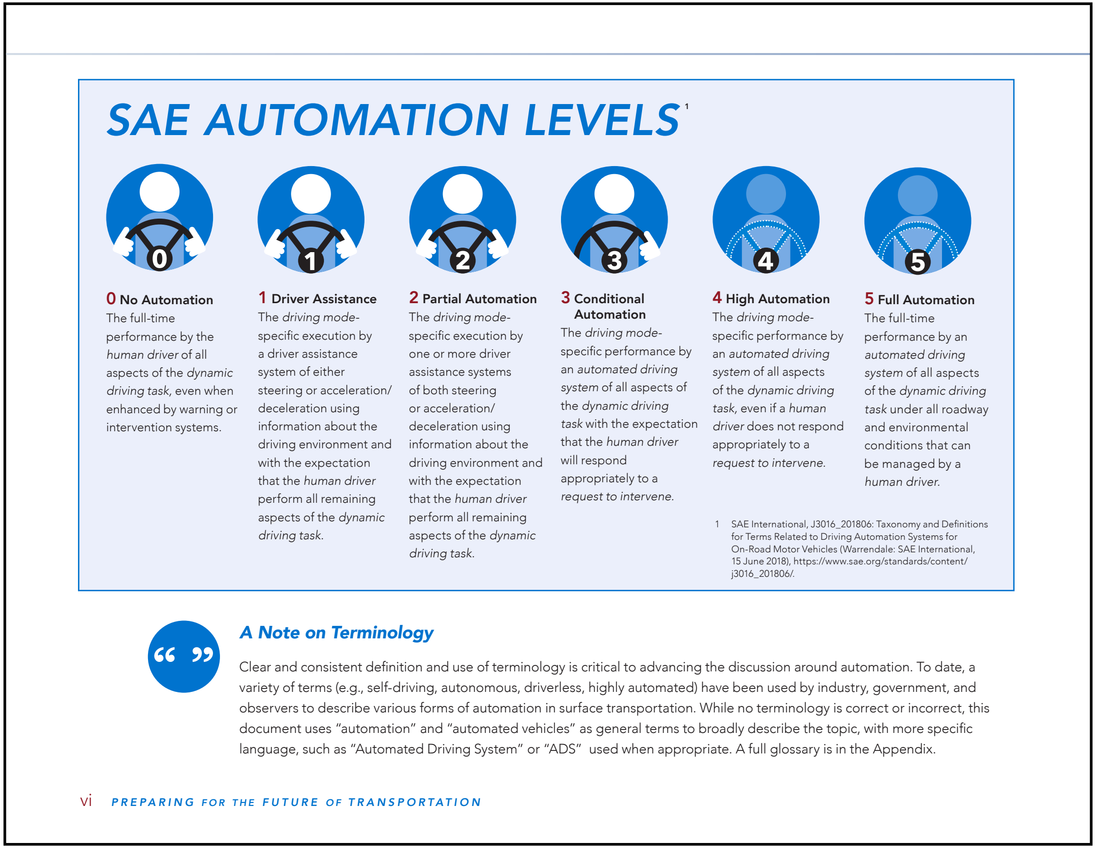
5. We will prepare proactively for automation.

U.S. DOT will provide guidance, best practices, pilot programs, and other assistance to help our partners plan and make the investments needed for a dynamic and flexible automated future. The Department also will prepare for complementary technologies that enhance the benefits of automation, such as communications between vehicles and the surrounding environment, but will not assume universal implementation of any particular approach.

6. We will protect and enhance the freedoms enjoyed by Americans.

U.S. DOT embraces the freedom of the open road, which includes the freedom for Americans to drive their own vehicles. We envision an environment in which automated vehicles operate alongside conventional, manually-driven vehicles and other road users. We will protect the ability of consumers to make the mobility choices that best suit their needs. We will support automation technologies that enhance individual freedom by expanding access to safe and independent mobility to people with disabilities and older Americans.

PREPARING FOR THE FUTURE OF TRANSPORTATION v



SAE AUTOMATION LEVELS[1]

0 No Automation
The full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems.

1 Driver Assistance
The *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*.

2 Partial Automation
The *driving mode*-specific execution by one or more driver assistance systems of both steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*.

3 Conditional Automation
The *driving mode*-specific performance by an *automated driving system* of all aspects of the *dynamic driving task* with the expectation that the *human driver* will respond appropriately to a *request to intervene*.

4 High Automation
The *driving mode*-specific performance by an *automated driving system* of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*.

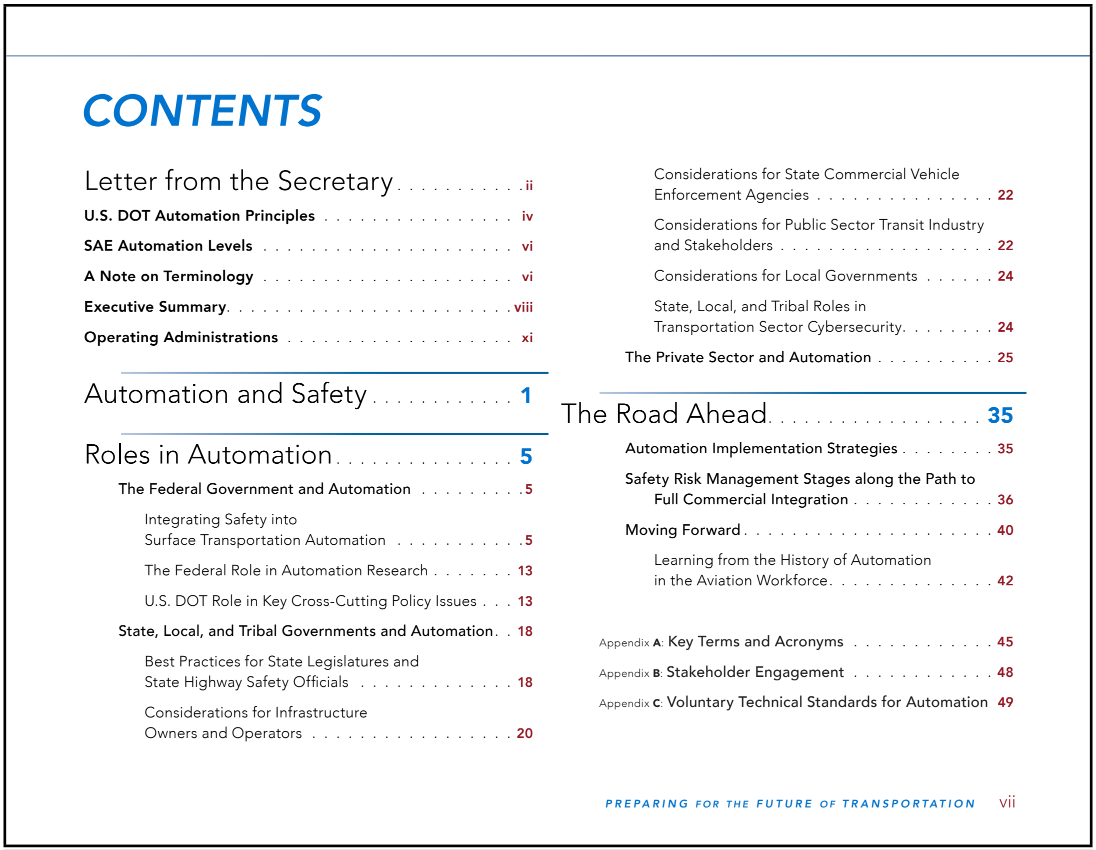
5 Full Automation
The full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*.

[1] SAE International, J3016_201806: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Warrendale: SAE International, 15 June 2018), https://www.sae.org/standards/content/j3016_201806/.

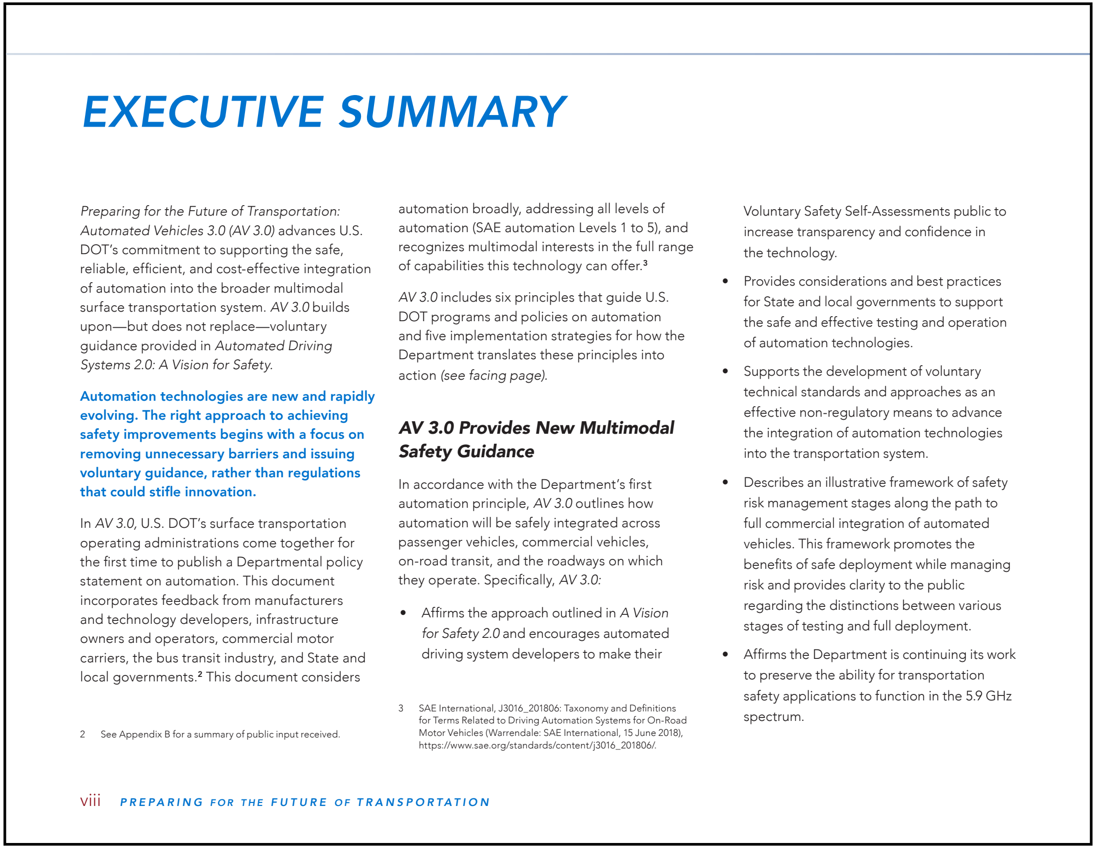
A Note on Terminology

Clear and consistent definition and use of terminology is critical to advancing the discussion around automation. To date, a variety of terms (e.g., self-driving, autonomous, driverless, highly automated) have been used by industry, government, and observers to describe various forms of automation in surface transportation. While no terminology is correct or incorrect, this document uses “automation” and “automated vehicles” as general terms to broadly describe the topic, with more specific language, such as “Automated Driving System” or “ADS” used when appropriate. A full glossary is in the Appendix.

vi PREPARING FOR THE FUTURE OF TRANSPORTATION



CONTENTS



EXECUTIVE SUMMARY

Preparing for the Future of Transportation: Automated Vehicles 3.0 (AV 3.0) advances U.S. DOT's commitment to supporting the safe, reliable, efficient, and cost-effective integration of automation into the broader multimodal surface transportation system. *AV 3.0* builds upon—but does not replace—voluntary guidance provided in *Automated Driving Systems 2.0: A Vision for Safety.*

Automation technologies are new and rapidly evolving. The right approach to achieving safety improvements begins with a focus on removing unnecessary barriers and issuing voluntary guidance, rather than regulations that could stifle innovation.

In *AV 3.0,* U.S. DOT's surface transportation operating administrations come together for the first time to publish a Departmental policy statement on automation. This document incorporates feedback from manufacturers and technology developers, infrastructure owners and operators, commercial motor carriers, the bus transit industry, and State and local governments.[2] This document considers automation broadly, addressing all levels of automation (SAE automation Levels 1 to 5), and recognizes multimodal interests in the full range of capabilities this technology can offer.[3]

AV 3.0 includes six principles that guide U.S. DOT programs and policies on automation and five implementation strategies for how the Department translates these principles into action *(see facing page).*

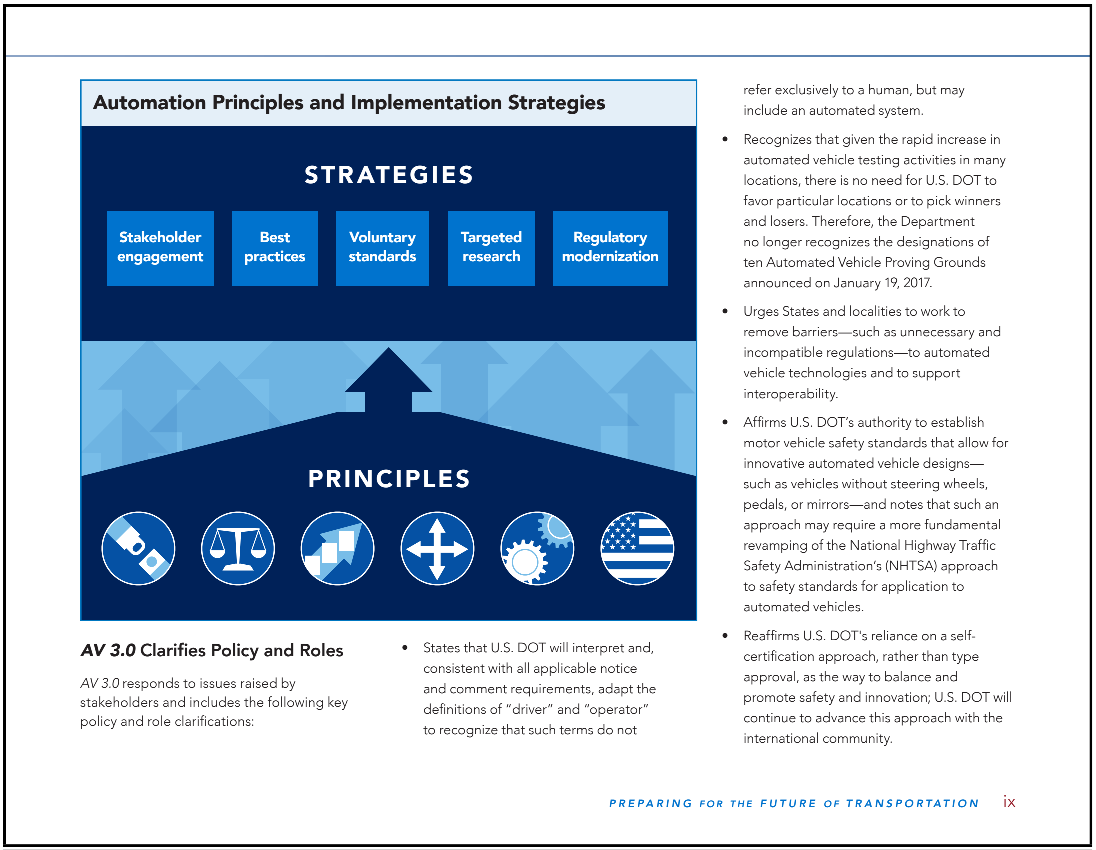
AV 3.0 Provides New Multimodal Safety Guidance

In accordance with the Department's first automation principle, *AV 3.0* outlines how automation will be safely integrated across passenger vehicles, commercial vehicles, on-road transit, and the roadways on which they operate. Specifically, *AV 3.0:*

- Affirms the approach outlined in *A Vision for Safety 2.0* and encourages automated driving system developers to make their Voluntary Safety Self-Assessments public to increase transparency and confidence in the technology.
- Provides considerations and best practices for State and local governments to support the safe and effective testing and operation of automation technologies.
- Supports the development of voluntary technical standards and approaches as an effective non-regulatory means to advance the integration of automation technologies into the transportation system.
- Describes an illustrative framework of safety risk management stages along the path to full commercial integration of automated vehicles. This framework promotes the benefits of safe deployment while managing risk and provides clarity to the public regarding the distinctions between various stages of testing and full deployment.
- Affirms the Department is continuing its work to preserve the ability for transportation safety applications to function in the 5.9 GHz spectrum.

[2] See Appendix B for a summary of public input received.

[3] SAE International, J3016_201806: Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Warrendale: SAE International, 15 June 2018), https://www.sae.org/standards/content/j3016_201806/.



Automation Principles and Implementation Strategies

STRATEGIES

Stakeholder engagement

Best practices

Voluntary standards

Targeted research

Regulatory modernization

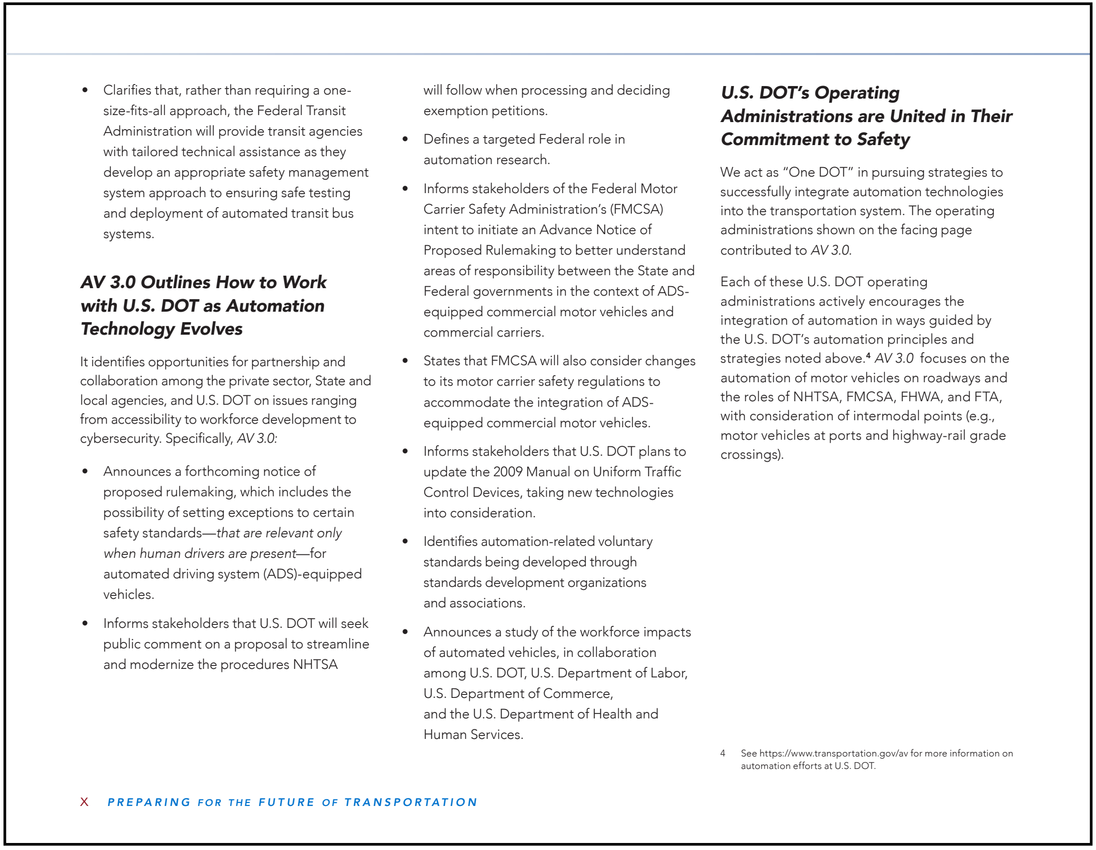
PRINCIPLES

AV 3.0 Clarifies Policy and Roles

AV 3.0 responds to issues raised by stakeholders and includes the following key policy and role clarifications:

- States that U.S. DOT will interpret and, consistent with all applicable notice and comment requirements, adapt the definitions of "driver" and "operator" to recognize that such terms do not refer exclusively to a human, but may include an automated system.
- Recognizes that given the rapid increase in automated vehicle testing activities in many locations, there is no need for U.S. DOT to favor particular locations or to pick winners and losers. Therefore, the Department no longer recognizes the designations of ten Automated Vehicle Proving Grounds announced on January 19, 2017.
- Urges States and localities to work to remove barriers—such as unnecessary and incompatible regulations—to automated vehicle technologies and to support interoperability.
- Affirms U.S. DOT's authority to establish motor vehicle safety standards that allow for innovative automated vehicle designs—such as vehicles without steering wheels, pedals, or mirrors—and notes that such an approach may require a more fundamental revamping of the National Highway Traffic Safety Administration's (NHTSA) approach to safety standards for application to automated vehicles.
- Reaffirms U.S. DOT's reliance on a self-certification approach, rather than type approval, as the way to balance and promote safety and innovation; U.S. DOT will continue to advance this approach with the international community.

PREPARING FOR THE FUTURE OF TRANSPORTATION ix



- Clarifies that, rather than requiring a one-size-fits-all approach, the Federal Transit Administration will provide transit agencies with tailored technical assistance as they develop an appropriate safety management system approach to ensuring safe testing and deployment of automated transit bus systems.

AV 3.0 Outlines How to Work with U.S. DOT as Automation Technology Evolves

It identifies opportunities for partnership and collaboration among the private sector, State and local agencies, and U.S. DOT on issues ranging from accessibility to workforce development to cybersecurity. Specifically, *AV 3.0:*

- Announces a forthcoming notice of proposed rulemaking, which includes the possibility of setting exceptions to certain safety standards—*that are relevant only when human drivers are present*—for automated driving system (ADS)-equipped vehicles.
- Informs stakeholders that U.S. DOT will seek public comment on a proposal to streamline and modernize the procedures NHTSA will follow when processing and deciding exemption petitions.
- Defines a targeted Federal role in automation research.
- Informs stakeholders of the Federal Motor Carrier Safety Administration's (FMCSA) intent to initiate an Advance Notice of Proposed Rulemaking to better understand areas of responsibility between the State and Federal governments in the context of ADS-equipped commercial motor vehicles and commercial carriers.
- States that FMCSA will also consider changes to its motor carrier safety regulations to accommodate the integration of ADS-equipped commercial motor vehicles.
- Informs stakeholders that U.S. DOT plans to update the 2009 Manual on Uniform Traffic Control Devices, taking new technologies into consideration.
- Identifies automation-related voluntary standards being developed through standards development organizations and associations.
- Announces a study of the workforce impacts of automated vehicles, in collaboration among U.S. DOT, U.S. Department of Labor, U.S. Department of Commerce, and the U.S. Department of Health and Human Services.

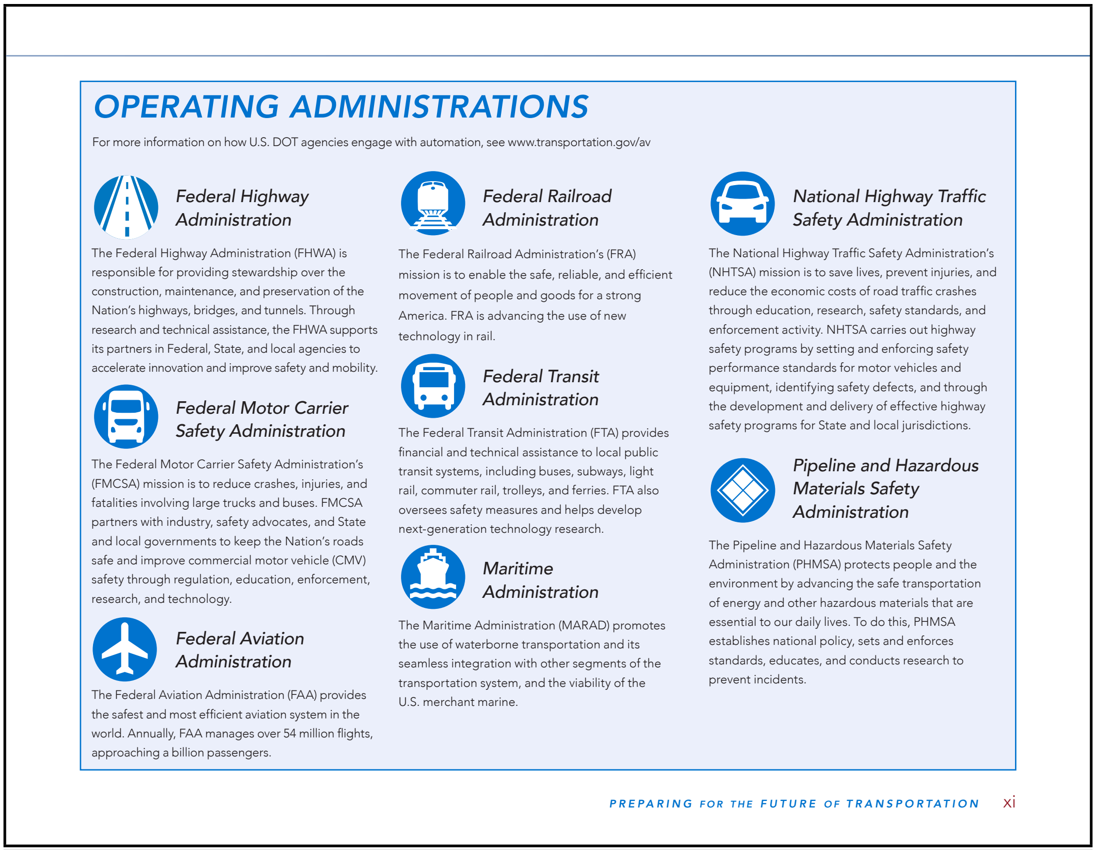
U.S. DOT's Operating Administrations are United in Their Commitment to Safety

We act as "One DOT" in pursuing strategies to successfully integrate automation technologies into the transportation system. The operating administrations shown on the facing page contributed to *AV 3.0.*

Each of these U.S. DOT operating administrations actively encourages the integration of automation in ways guided by the U.S. DOT's automation principles and strategies noted above.[4] *AV 3.0* focuses on the automation of motor vehicles on roadways and the roles of NHTSA, FMCSA, FHWA, and FTA, with consideration of intermodal points (e.g., motor vehicles at ports and highway-rail grade crossings).

[4] See https://www.transportation.gov/av for more information on automation efforts at U.S. DOT.

x PREPARING FOR THE FUTURE OF TRANSPORTATION



OPERATING ADMINISTRATIONS

For more information on how U.S. DOT agencies engage with automation, see www.transportation.gov/av

Federal Highway Administration

The Federal Highway Administration (FHWA) is responsible for providing stewardship over the construction, maintenance, and preservation of the Nation's highways, bridges, and tunnels. Through research and technical assistance, the FHWA supports its partners in Federal, State, and local agencies to accelerate innovation and improve safety and mobility.

Federal Motor Carrier Safety Administration

The Federal Motor Carrier Safety Administration's (FMCSA) mission is to reduce crashes, injuries, and fatalities involving large trucks and buses. FMCSA partners with industry, safety advocates, and State and local governments to keep the Nation's roads safe and improve commercial motor vehicle (CMV) safety through regulation, education, enforcement, research, and technology.

Federal Aviation Administration

The Federal Aviation Administration (FAA) provides the safest and most efficient aviation system in the world. Annually, FAA manages over 54 million flights, approaching a billion passengers.

Federal Railroad Administration

The Federal Railroad Administration's (FRA) mission is to enable the safe, reliable, and efficient movement of people and goods for a strong America. FRA is advancing the use of new technology in rail.

Federal Transit Administration

The Federal Transit Administration (FTA) provides financial and technical assistance to local public transit systems, including buses, subways, light rail, commuter rail, trolleys, and ferries. FTA also oversees safety measures and helps develop next-generation technology research.

Maritime Administration

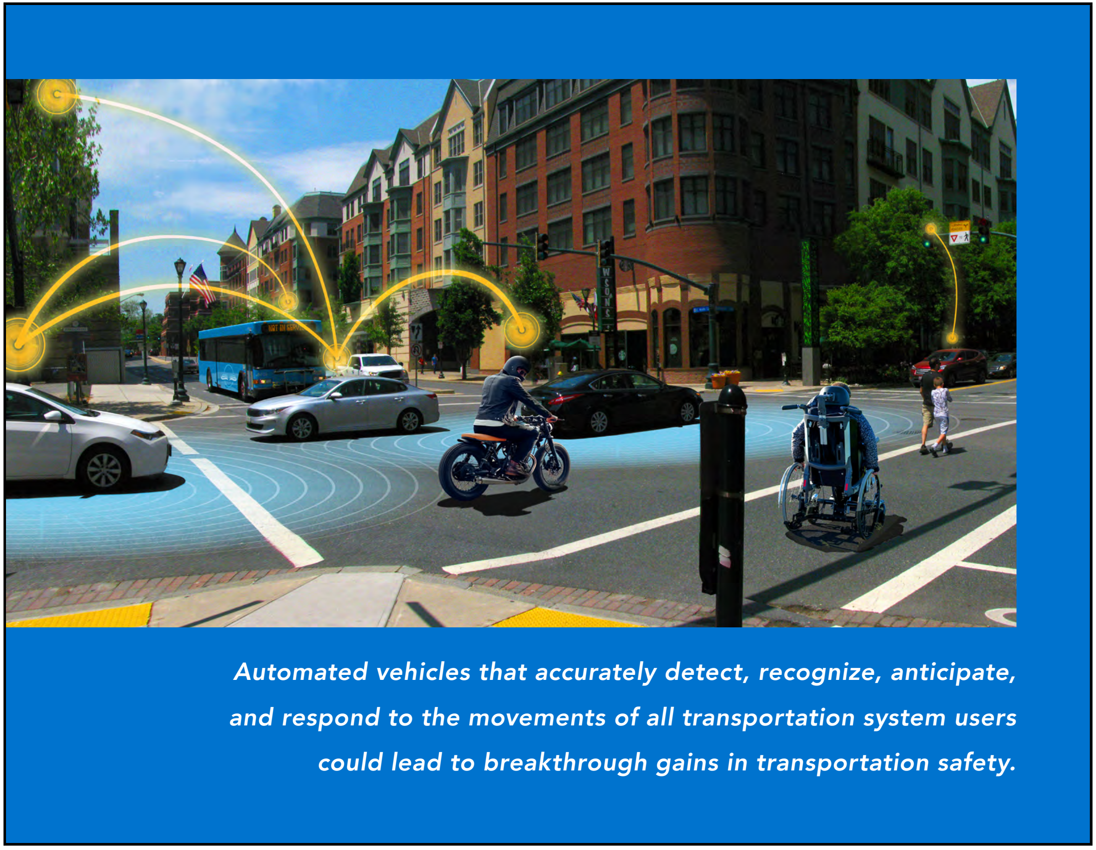
The Maritime Administration (MARAD) promotes the use of waterborne transportation and its seamless integration with other segments of the transportation system, and the viability of the U.S. merchant marine.

National Highway Traffic Safety Administration

The National Highway Traffic Safety Administration's (NHTSA) mission is to save lives, prevent injuries, and reduce the economic costs of road traffic crashes through education, research, safety standards, and enforcement activity. NHTSA carries out highway safety programs by setting and enforcing safety performance standards for motor vehicles and equipment, identifying safety defects, and through the development and delivery of effective highway safety programs for State and local jurisdictions.

Pipeline and Hazardous Materials Safety Administration

The Pipeline and Hazardous Materials Safety Administration (PHMSA) protects people and the environment by advancing the safe transportation of energy and other hazardous materials that are essential to our daily lives. To do this, PHMSA establishes national policy, sets and enforces standards, educates, and conducts research to prevent incidents.



Automated vehicles that accurately detect, recognize, anticipate, and respond to the movements of all transportation system users could lead to breakthrough gains in transportation safety.

INTRODUCTION: AUTOMATION AND SAFETY

The United States surface transportation system provides tremendous mobility benefits, including widespread access to jobs, goods, and services. It also connects many remote regions of the country to the larger economy. These benefits, however, come with significant safety challenges, as motor vehicle crashes remain a leading cause of death, with an estimated 37,133 lives lost on U.S. roads in 2017. Traditional safety programs and policies have made road travel significantly safer than in the past, but there is much room to improve traffic fatality and injury rates.

Automated vehicles that accurately detect, recognize, anticipate, and respond to the movements of all transportation system users could lead to breakthrough gains in transportation safety. Unlike human drivers, automation technologies are not prone to distraction, fatigue, or impaired driving, which contribute to a significant portion of surface transportation fatalities. Automated vehicle technologies that are carefully integrated into motor vehicles could help vehicle operators detect and avoid bicyclists, motorcyclists, pedestrians, and other vulnerable users on our roadways, and increase safety across the surface transportation system. Their potential to reduce deaths and injuries on the Nation's roadways cannot be overstated.

Automated vehicles rely on sensors and software that allow an expansive view of the environment across a range of lighting and weather conditions. They can quickly learn and adapt to new driving situations by learning from previous experience through software updates. Fully realizing the life-saving potential of automated vehicles, however, will require careful risk management as new technologies are introduced and adopted across the surface transportation system.

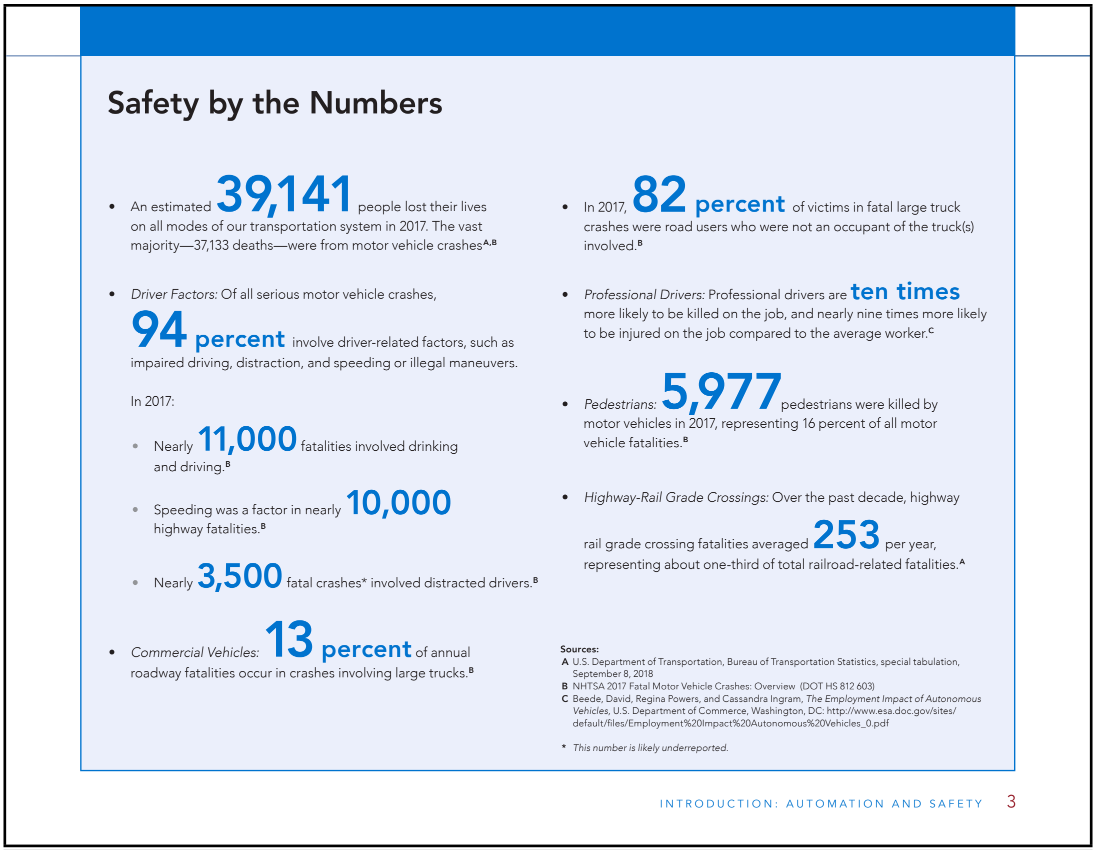
To support the deployment of safe automation technologies, the Department released *A Vision for Safety 2.0* in September 2017, which included 12 automated driving system (ADS) safety elements to help industry partners analyze, identify, and resolve safety considerations using best practices—all before deployment. The voluntary guidance outlined in *A Vision for Safety 2.0* on the design, testing, and safe deployment of ADS remains central to U.S. DOT's approach. ADS developers are encouraged to use these safety elements to publish safety self-assessments to describe to the public how they are identifying and addressing potential safety issues.

On-road testing and early deployments are important to improving automated vehicle performance and allowing them to reach their full performance potential. Careful real-world testing allows developers to identify and rapidly fix system shortcomings, not just on individual vehicles but across fleets. Reasonable risks must be addressed through the application of robust systems engineering processes, testing protocols, and functional safety best practices, such as those documented in *A Vision for Safety 2.0*.[5] However, delaying or unduly hampering automated vehicle testing until all specific risks have been identified and eliminated means delaying the realization of global reductions in risk.

AV 3.0 maintains U.S. DOT's primary focus on safety, while expanding the discussion to other aspects and modes of surface transportation. *AV 3.0* introduces a comprehensive, multimodal approach toward safely integrating automation.

AV 3.0 *introduces a comprehensive, multimodal approach toward safely integrating automation.*

[5] As documented in *A Vision for Safety 2.0*, ADS developers should consider employing systems engineering guidance, best practices, design principles, and standards developed by established and accredited standards-developing organizations (as applicable) such as the International Standards Organization (ISO) and SAE International as well as standards and processes available from other industries, such as aviation, space, and the military and other applicable standards or internal company processes as they are relevant and applicable. They should also consider available and emerging approaches to risk mitigation, such as methodologies that focus on functional safety (e.g., ISO 26262) and safety of the intended functionality.



Safety by the Numbers

- An estimated **39,141** people lost their lives on all modes of our transportation system in 2017. The vast majority—37,133 deaths—were from motor vehicle crashes[A,B]

- *Driver Factors:* Of all serious motor vehicle crashes, **94 percent** involve driver-related factors, such as impaired driving, distraction, and speeding or illegal maneuvers.

  In 2017:

  - Nearly **11,000** fatalities involved drinking and driving.[B]
  - Speeding was a factor in nearly **10,000** highway fatalities.[B]
  - Nearly **3,500** fatal crashes* involved distracted drivers.[B]

- *Commercial Vehicles:* **13 percent** of annual roadway fatalities occur in crashes involving large trucks.[B]

- In 2017, **82 percent** of victims in fatal large truck crashes were road users who were not an occupant of the truck(s) involved.[B]

- *Professional Drivers:* Professional drivers are **ten times** more likely to be killed on the job, and nearly nine times more likely to be injured on the job compared to the average worker.[C]

- *Pedestrians:* **5,977** pedestrians were killed by motor vehicles in 2017, representing 16 percent of all motor vehicle fatalities.[B]

- *Highway-Rail Grade Crossings:* Over the past decade, highway rail grade crossing fatalities averaged **253** per year, representing about one-third of total railroad-related fatalities.[A]

Sources:

A U.S. Department of Transportation, Bureau of Transportation Statistics, special tabulation, September 8, 2018

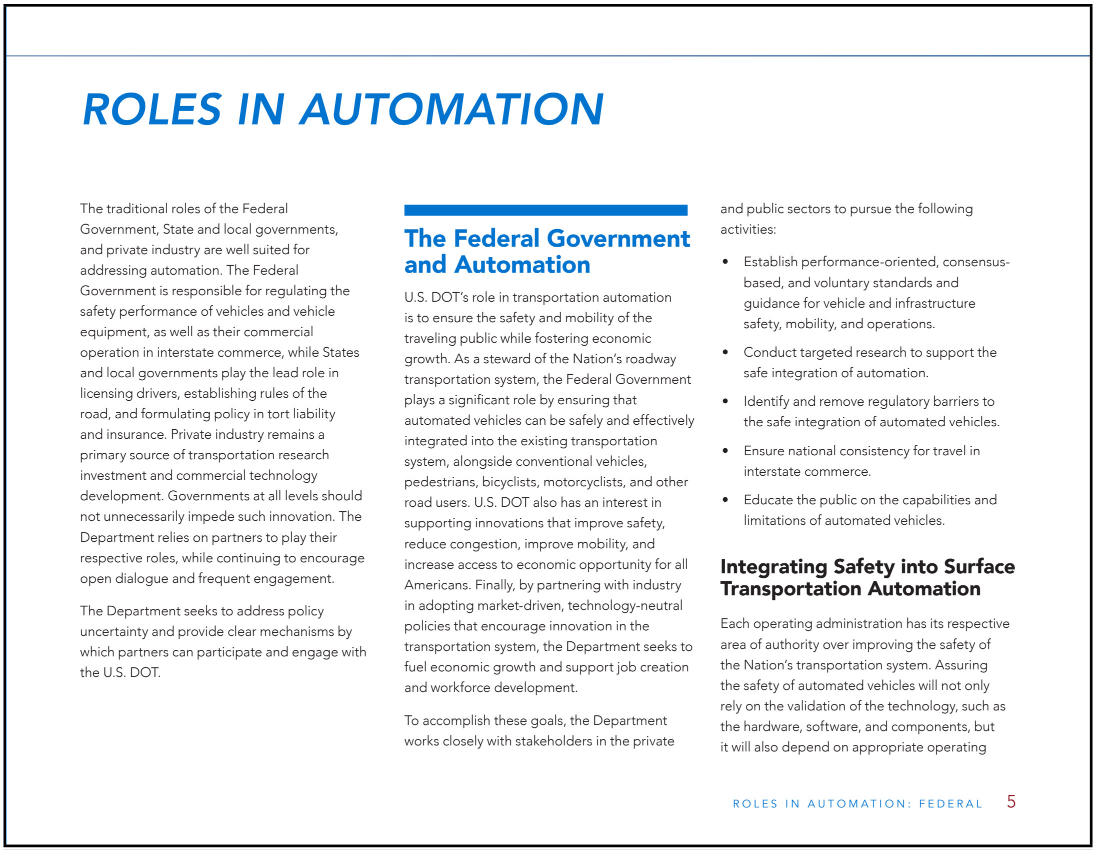
B NHTSA 2017 Fatal Motor Vehicle Crashes: Overview (DOT HS 812 603)

C Beede, David, Regina Powers, and Cassandra Ingram, *The Employment Impact of Autonomous Vehicles*, U.S. Department of Commerce, Washington, DC: http://www.esa.doc.gov/sites/default/files/Employment%20Impact%20Autonomous%20Vehicles_0.pdf

* *This number is likely underreported.*

INTRODUCTION: AUTOMATION AND SAFETY 3

Only by working in partnership can the public and the private sector improve the safety, security, and accessibility of automation technologies and address the concerns of the general public.



ROLES IN AUTOMATION

The traditional roles of the Federal Government, State and local governments, and private industry are well suited for addressing automation. The Federal Government is responsible for regulating the safety performance of vehicles and vehicle equipment, as well as their commercial operation in interstate commerce, while States and local governments play the lead role in licensing drivers, establishing rules of the road, and formulating policy in tort liability and insurance. Private industry remains a primary source of transportation research investment and commercial technology development. Governments at all levels should not unnecessarily impede such innovation. The Department relies on partners to play their respective roles, while continuing to encourage open dialogue and frequent engagement.

The Department seeks to address policy uncertainty and provide clear mechanisms by which partners can participate and engage with the U.S. DOT.

The Federal Government and Automation

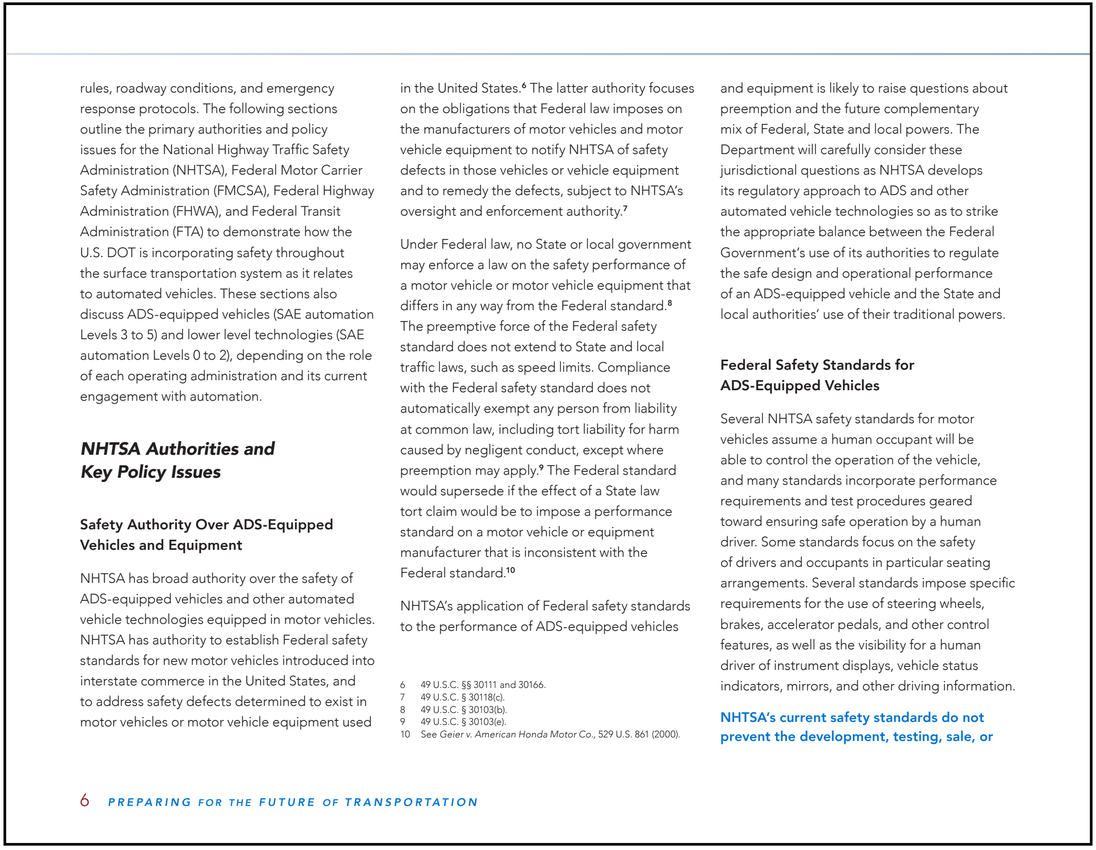
U.S. DOT's role in transportation automation is to ensure the safety and mobility of the traveling public while fostering economic growth. As a steward of the Nation's roadway transportation system, the Federal Government plays a significant role by ensuring that automated vehicles can be safely and effectively integrated into the existing transportation system, alongside conventional vehicles, pedestrians, bicyclists, motorcyclists, and other road users. U.S. DOT also has an interest in supporting innovations that improve safety, reduce congestion, improve mobility, and increase access to economic opportunity for all Americans. Finally, by partnering with industry in adopting market-driven, technology-neutral policies that encourage innovation in the transportation system, the Department seeks to fuel economic growth and support job creation and workforce development.

To accomplish these goals, the Department works closely with stakeholders in the private and public sectors to pursue the following activities:

- Establish performance-oriented, consensus-based, and voluntary standards and guidance for vehicle and infrastructure safety, mobility, and operations.
- Conduct targeted research to support the safe integration of automation.
- Identify and remove regulatory barriers to the safe integration of automated vehicles.
- Ensure national consistency for travel in interstate commerce.
- Educate the public on the capabilities and limitations of automated vehicles.

Integrating Safety into Surface Transportation Automation

Each operating administration has its respective area of authority over improving the safety of the Nation's transportation system. Assuring the safety of automated vehicles will not only rely on the validation of the technology, such as the hardware, software, and components, but it will also depend on appropriate operating



rules, roadway conditions, and emergency response protocols. The following sections outline the primary authorities and policy issues for the National Highway Traffic Safety Administration (NHTSA), Federal Motor Carrier Safety Administration (FMCSA), Federal Highway Administration (FHWA), and Federal Transit Administration (FTA) to demonstrate how the U.S. DOT is incorporating safety throughout the surface transportation system as it relates to automated vehicles. These sections also discuss ADS-equipped vehicles (SAE automation Levels 3 to 5) and lower level technologies (SAE automation Levels 0 to 2), depending on the role of each operating administration and its current engagement with automation.

NHTSA Authorities and Key Policy Issues

Safety Authority Over ADS-Equipped Vehicles and Equipment

NHTSA has broad authority over the safety of ADS-equipped vehicles and other automated vehicle technologies equipped in motor vehicles. NHTSA has authority to establish Federal safety standards for new motor vehicles introduced into interstate commerce in the United States, and to address safety defects determined to exist in motor vehicles or motor vehicle equipment used in the United States.[6] The latter authority focuses on the obligations that Federal law imposes on the manufacturers of motor vehicles and motor vehicle equipment to notify NHTSA of safety defects in those vehicles or vehicle equipment and to remedy the defects, subject to NHTSA's oversight and enforcement authority.[7]

Under Federal law, no State or local government may enforce a law on the safety performance of a motor vehicle or motor vehicle equipment that differs in any way from the Federal standard.[8] The preemptive force of the Federal safety standard does not extend to State and local traffic laws, such as speed limits. Compliance with the Federal safety standard does not automatically exempt any person from liability at common law, including tort liability for harm caused by negligent conduct, except where preemption may apply.[9] The Federal standard would supersede if the effect of a State law tort claim would be to impose a performance standard on a motor vehicle or equipment manufacturer that is inconsistent with the Federal standard.[10]

NHTSA's application of Federal safety standards to the performance of ADS-equipped vehicles and equipment is likely to raise questions about preemption and the future complementary mix of Federal, State and local powers. The Department will carefully consider these jurisdictional questions as NHTSA develops its regulatory approach to ADS and other automated vehicle technologies so as to strike the appropriate balance between the Federal Government's use of its authorities to regulate the safe design and operational performance of an ADS-equipped vehicle and the State and local authorities' use of their traditional powers.

Federal Safety Standards for ADS-Equipped Vehicles

Several NHTSA safety standards for motor vehicles assume a human occupant will be able to control the operation of the vehicle, and many standards incorporate performance requirements and test procedures geared toward ensuring safe operation by a human driver. Some standards focus on the safety of drivers and occupants in particular seating arrangements. Several standards impose specific requirements for the use of steering wheels, brakes, accelerator pedals, and other control features, as well as the visibility for a human driver of instrument displays, vehicle status indicators, mirrors, and other driving information.

NHTSA's current safety standards do not prevent the development, testing, sale, or

[6] 49 U.S.C. §§ 30111 and 30166.
[7] 49 U.S.C. § 30118(c).
[8] 49 U.S.C. § 30103(b).
[9] 49 U.S.C. § 30103(e).
[10] See *Geier v. American Honda Motor Co.*, 529 U.S. 861 (2000).

use of ADS built into vehicles that maintain the traditional cabin and control features of human-operated vehicles. However, some Level 4 and 5 automated vehicles may be designed to be controlled entirely by an ADS, and the interior of the vehicle may be configured without human controls. There may be no steering wheel, accelerator pedal, brakes, mirrors, or information displays for human use. For such ADS-equipped vehicles, NHTSA's current safety standards constitute an unintended regulatory barrier to innovation.

The Department, through NHTSA, intends to reconsider the necessity and appropriateness of its current safety standards as applied to ADS-equipped vehicles. **In an upcoming rulemaking, NHTSA plans to seek comment on proposed changes to particular safety standards to accommodate automated vehicle technologies and the possibility of setting exceptions to certain standards—*that are relevant only when human drivers are present*—for ADS-equipped vehicles.**

Going forward, NHTSA may also consider a more fundamental revamping of its approach to safety standards for application to automated vehicles. However, reliance on a self-certification approach, instead of type approval, more appropriately balances and promotes safety and innovation; U.S. DOT will continue to advance this approach with the international community. NHTSA's current statutory authority to establish motor vehicle safety standards is sufficiently flexible to accommodate the design and performance of different ADS concepts in new vehicle configurations.

NHTSA recognizes that the accelerating pace of technological change, especially in the development of software used in ADS-equipped vehicles, requires a new approach to the formulation of the Federal Motor Vehicle Safety Standards (FMVSS). The pace of innovation in automated vehicle technologies is incompatible with lengthy rulemaking proceedings and highly prescriptive and feature-specific or design-specific safety standards. Future motor vehicle safety standards will need to be more flexible and responsive, technology-neutral, and performance-oriented to accommodate rapid technological innovation. They may incorporate simpler and more general requirements designed to validate that an ADS can safely navigate the real-world roadway environment, including unpredictable hazards, obstacles, and interactions with other vehicles and pedestrians who may not always adhere to the traffic laws or follow expected patterns of behavior. Existing standards assume that a vehicle may be driven anywhere, but future standards will need to take into account that the operational design domain (ODD) for a particular ADS within a vehicle is likely to be limited in some ways that may be unique to that system. For example, not all Level 3 vehicles will have the same ODD.

Performance-based safety standards could require manufacturers to use test methods, such as sophisticated obstacle-course-based test regimes, sufficient to validate that their ADS-equipped vehicles can reliably handle the normal range of everyday driving scenarios as well as unusual and unpredictable scenarios. Standards could be designed to account for factors such as variations in weather, traffic, and roadway conditions within a given system's ODD, as well as sudden and unpredictable actions by other road users. Test procedures could also be developed to ensure that an ADS does not operate outside of the ODD established by the manufacturer. Standards could provide for a range of potential behaviors—e.g., speed, distance, angles, and size—for surrogate vehicles, pedestrians, and other obstacles that ADS-equipped vehicles would need to detect and avoid. Other approaches, such as computer simulation and requirements expressed in terms of mathematical functions could be considered, as Federal law does not require that NHTSA's safety standards rely on physical tests and measurements, only that they be objective, repeatable, and transparent.

Exemptions from FMVSS for ADS Purposes

NHTSA values a streamlined and modernized exemptions procedure, and removing unnecessary delays. **NHTSA intends to seek**

public comment on a proposal to streamline and modernize procedures the Agency will follow when processing and deciding exemption petitions. Among other things, the proposed changes will remove unnecessary delays in seeking public comment as part of the exemption process, and clarify and update the types of information needed to support such petitions. The statutory provision authorizing NHTSA to grant exemptions from FMVSS provides sufficient flexibility to accommodate a wide array of automated operations, particularly for manufacturers seeking to engage in research, testing, and demonstration projects.[11]

FMCSA Authorities and Key Policy Issues

Safety Authority Over Commercial Motor Vehicle Operations, Drivers, and Maintenance

The Department, through FMCSA, regulates the safety of commercial motor carriers operating in interstate commerce, the qualifications and safety of commercial motor vehicle drivers, and the safe operation of commercial trucks and motor coaches.[12] The best way to accomplish FMCSA's core mission of reducing fatalities and crashes involving large trucks and buses is to avoid unnecessary barriers to the development of ADS in commercial vehicles.

As automation introduces new policy questions, FMCSA will work with (1) industry, State governments, and other partners to further the safe operation of ADS-equipped commercial vehicles, and (2) law enforcement, inspection officers, and first responders to create new techniques and protocols.

In order to develop experience with the technology, demonstrate its capabilities, and socialize the idea of automated vehicles on the road with traditional vehicles, FMCSA will continue to hold public demonstrations of the technology—such as the recent truck platooning demonstration on the I-66 Corridor co-hosted with FHWA—with key stakeholders such as law enforcement.

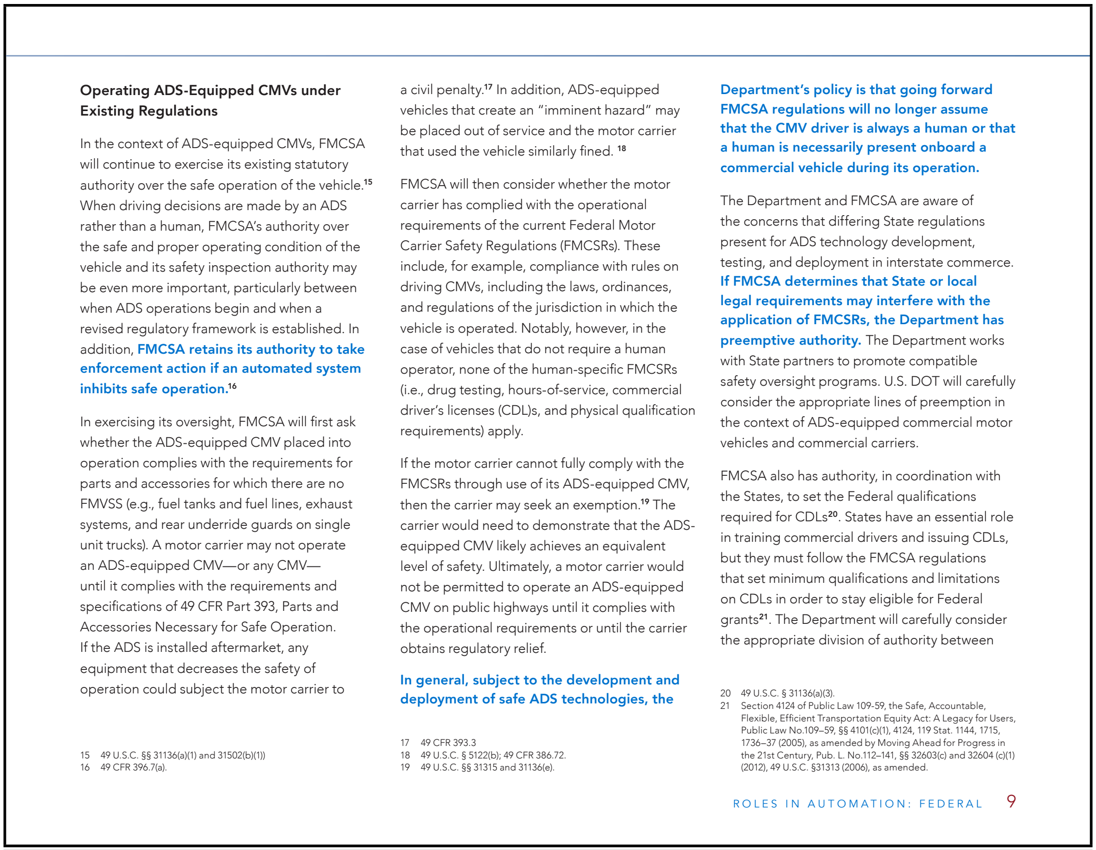
FMCSA consults with NHTSA on matters related to motor carrier safety.[13] NHTSA and FMCSA have different but complementary authorities over the safety of commercial motor vehicles (CMVs) and commercial vehicle equipment. NHTSA has exclusive authority to prescribe Federal safety standards for new motor vehicles, including trucks and motor coaches, and oversees actions that manufacturers take to remedy known safety defects in motor vehicles and motor vehicle equipment.[14] NHTSA and FMCSA collaborate and consult to develop and enforce safety requirements that apply to the operation and maintenance of vehicles by existing commercial motor carriers. They will continue to do so in the context of ADS-equipped commercial motor vehicles. FMCSA also works closely with States and private stakeholders to develop and enforce safety standards related to the inspection, maintenance, and repair of commercial motor vehicles.

[11] 49 U.S.C. § 30114

[12] 49 U.S.C. § 31502; 49 U.S.C. chapter 311, subchapter III; 49 U.S.C. chapter 313. Additional statutory authority includes the Hazardous Materials Transportation Uniform Safety Act of 1990 (Pub. L. 101-615, 104 Stat. 3244), codified at 49 U.S.C. Chapter 51; and the ICC Termination Act of 1995 (Pub. L. 104-88, 109 Stat. 803), codified at 49 U.S.C. Chapters 135-149. Note that FMCSA's statutory authority also authorizes the Agency's enforcement of the Hazardous Materials Regulations (HMRs) and the Federal Motor Carrier Commercial Regulations (FMCCRs),. 49 U.S.C. chapter 311, subchapters I and III; chapter 313; and section 31502

[13] 49 U.S.C. § 113(i).

[14] See 49 U.S.C. §§ 30111 and 30166



Operating ADS-Equipped CMVs under Existing Regulations

In the context of ADS-equipped CMVs, FMCSA will continue to exercise its existing statutory authority over the safe operation of the vehicle.[15] When driving decisions are made by an ADS rather than a human, FMCSA's authority over the safe and proper operating condition of the vehicle and its safety inspection authority may be even more important, particularly between when ADS operations begin and when a revised regulatory framework is established. In addition, **FMCSA retains its authority to take enforcement action if an automated system inhibits safe operation.**[16]

In exercising its oversight, FMCSA will first ask whether the ADS-equipped CMV placed into operation complies with the requirements for parts and accessories for which there are no FMVSS (e.g., fuel tanks and fuel lines, exhaust systems, and rear underride guards on single unit trucks). A motor carrier may not operate an ADS-equipped CMV—or any CMV—until it complies with the requirements and specifications of 49 CFR Part 393, Parts and Accessories Necessary for Safe Operation. If the ADS is installed aftermarket, any equipment that decreases the safety of operation could subject the motor carrier to a civil penalty.[17] In addition, ADS-equipped vehicles that create an "imminent hazard" may be placed out of service and the motor carrier that used the vehicle similarly fined. [18]

FMCSA will then consider whether the motor carrier has complied with the operational requirements of the current Federal Motor Carrier Safety Regulations (FMCSRs). These include, for example, compliance with rules on driving CMVs, including the laws, ordinances, and regulations of the jurisdiction in which the vehicle is operated. Notably, however, in the case of vehicles that do not require a human operator, none of the human-specific FMCSRs (i.e., drug testing, hours-of-service, commercial driver's licenses (CDL)s, and physical qualification requirements) apply.

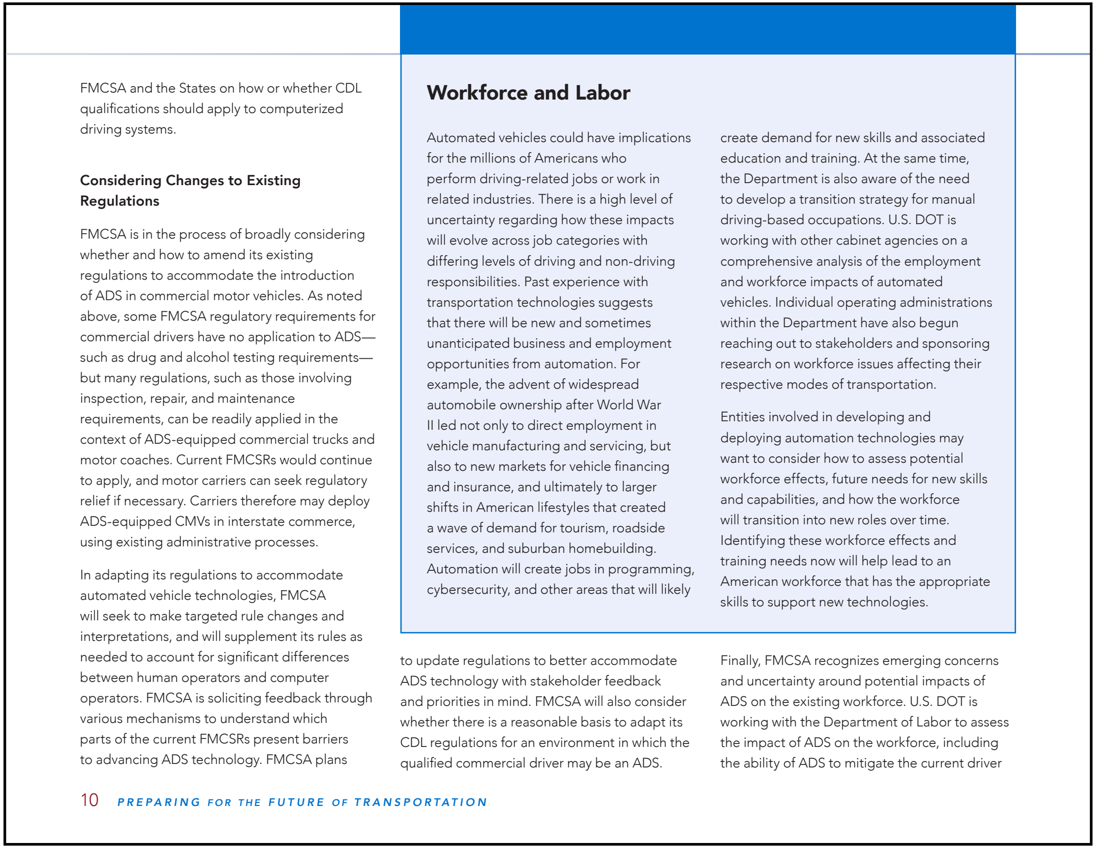
If the motor carrier cannot fully comply with the FMCSRs through use of its ADS-equipped CMV, then the carrier may seek an exemption.[19] The carrier would need to demonstrate that the ADS-equipped CMV likely achieves an equivalent level of safety. Ultimately, a motor carrier would not be permitted to operate an ADS-equipped CMV on public highways until it complies with the operational requirements or until the carrier obtains regulatory relief.

In general, subject to the development and deployment of safe ADS technologies, the Department's policy is that going forward FMCSA regulations will no longer assume that the CMV driver is always a human or that a human is necessarily present onboard a commercial vehicle during its operation.

The Department and FMCSA are aware of the concerns that differing State regulations present for ADS technology development, testing, and deployment in interstate commerce. **If FMCSA determines that State or local legal requirements may interfere with the application of FMCSRs, the Department has preemptive authority.** The Department works with State partners to promote compatible safety oversight programs. U.S. DOT will carefully consider the appropriate lines of preemption in the context of ADS-equipped commercial motor vehicles and commercial carriers.

FMCSA also has authority, in coordination with the States, to set the Federal qualifications required for CDLs[20]. States have an essential role in training commercial drivers and issuing CDLs, but they must follow the FMCSA regulations that set minimum qualifications and limitations on CDLs in order to stay eligible for Federal grants[21]. The Department will carefully consider the appropriate division of authority between

[15] 49 U.S.C. §§ 31136(a)(1) and 31502(b)(1))
[16] 49 CFR 396.7(a).
[17] 49 CFR 393.3
[18] 49 U.S.C. § 5122(b); 49 CFR 386.72.
[19] 49 U.S.C. §§ 31315 and 31136(e).
[20] 49 U.S.C. § 31136(a)(3).
[21] Section 4124 of Public Law 109-59, the Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users, Public Law No.109–59, §§ 4101(c)(1), 4124, 119 Stat. 1144, 1715, 1736–37 (2005), as amended by Moving Ahead for Progress in the 21st Century, Pub. L. No.112–141, §§ 32603(c) and 32604 (c)(1) (2012), 49 U.S.C. §31313 (2006), as amended.



FMCSA and the States on how or whether CDL qualifications should apply to computerized driving systems.

Considering Changes to Existing Regulations

FMCSA is in the process of broadly considering whether and how to amend its existing regulations to accommodate the introduction of ADS in commercial motor vehicles. As noted above, some FMCSA regulatory requirements for commercial drivers have no application to ADS—such as drug and alcohol testing requirements—but many regulations, such as those involving inspection, repair, and maintenance requirements, can be readily applied in the context of ADS-equipped commercial trucks and motor coaches. Current FMCSRs would continue to apply, and motor carriers can seek regulatory relief if necessary. Carriers therefore may deploy ADS-equipped CMVs in interstate commerce, using existing administrative processes.

In adapting its regulations to accommodate automated vehicle technologies, FMCSA will seek to make targeted rule changes and interpretations, and will supplement its rules as needed to account for significant differences between human operators and computer operators. FMCSA is soliciting feedback through various mechanisms to understand which parts of the current FMCSRs present barriers to advancing ADS technology. FMCSA plans to update regulations to better accommodate ADS technology with stakeholder feedback and priorities in mind. FMCSA will also consider whether there is a reasonable basis to adapt its CDL regulations for an environment in which the qualified commercial driver may be an ADS.

Finally, FMCSA recognizes emerging concerns and uncertainty around potential impacts of ADS on the existing workforce. U.S. DOT is working with the Department of Labor to assess the impact of ADS on the workforce, including the ability of ADS to mitigate the current driver

Workforce and Labor

Automated vehicles could have implications for the millions of Americans who perform driving-related jobs or work in related industries. There is a high level of uncertainty regarding how these impacts will evolve across job categories with differing levels of driving and non-driving responsibilities. Past experience with transportation technologies suggests that there will be new and sometimes unanticipated business and employment opportunities from automation. For example, the advent of widespread automobile ownership after World War II led not only to direct employment in vehicle manufacturing and servicing, but also to new markets for vehicle financing and insurance, and ultimately to larger shifts in American lifestyles that created a wave of demand for tourism, roadside services, and suburban homebuilding. Automation will create jobs in programming, cybersecurity, and other areas that will likely create demand for new skills and associated education and training. At the same time, the Department is also aware of the need to develop a transition strategy for manual driving-based occupations. U.S. DOT is working with other cabinet agencies on a comprehensive analysis of the employment and workforce impacts of automated vehicles. Individual operating administrations within the Department have also begun reaching out to stakeholders and sponsoring research on workforce issues affecting their respective modes of transportation.

Entities involved in developing and deploying automation technologies may want to consider how to assess potential workforce effects, future needs for new skills and capabilities, and how the workforce will transition into new roles over time. Identifying these workforce effects and training needs now will help lead to an American workforce that has the appropriate skills to support new technologies.

shortage in the motor carrier industry. The study will also look at longer-term needs for future workforce skills and at the demand for a transportation system that relies on ADS technology.

FHWA's Authorities Over Traffic Control Devices

U.S. DOT recognizes that the quality and uniformity of road markings, signage, and other traffic control devices support safe and efficient driving by both human drivers and automated vehicles.

As part of its role to support State and local governments in the design, construction, and maintenance of the Nation's roads, FHWA administers the Manual on Uniform Traffic Control Devices (MUTCD).[22] The MUTCD is recognized as the national standard for all traffic control devices installed on any street, highway, bikeway, or private road open to public travel. Traffic control devices generally refer to signs, signals, markings, and other devices used to regulate or guide traffic on a street, highway, and other facilities. FHWA, in partnership with key stakeholder associations and the practitioner community, is conducting research and device experimentation for overall improvements to the manual, and to better understand the specific needs of the emerging automated vehicle technologies. Incorporating existing interim approved devices, experimentations, and other identified proposed changes into the updated MUTCD will help humans and emerging automated vehicles to interpret the roadway. FHWA will use current research to supplement knowledge regarding different sensor and machine vision system capabilities relative to interpreting traffic control devices. **As part of this effort, FHWA will pursue an update to the 2009 MUTCD that will take into consideration these new technologies and other needs.**

FTA's Safety Authority Over Public Transportation

Safety issues are the highest priority for all providers of public transportation. In recent years, Congress has granted FTA significant new safety authorities that have expanded the Agency's role as a safety oversight regulatory body.[23] Consequently, FTA developed and published a National Public Transportation Safety Plan (NSP).[24] The NSP functions as FTA's strategic plan and primary guidance document for improving transit safety performance; a policy document and communications tool; and a repository of standards, guidance, best practices, tools, technical assistance, and other resources.

A key foundational component of FTA's safety authority is the new Public Transportation Agency Safety Plan (PTASP) rule.[25] The PTASP rule, which FTA issued on July 18, 2018, and which becomes effective on July 19, 2019, is applicable to transit agencies that operate rail fixed-guideway and/or bus services. Transit agencies must develop, certify, and implement an agency safety plan by July 20, 2020. The PTASP rule requires transit agencies to incorporate Safety Management System (SMS) policies and procedures as they develop their individual safety plans. The PTASP rule sets scalable and flexible requirements for public transportation agencies by requiring them to establish appropriate safety objectives; to identify safety risks and hazards and to develop plans to mitigate those risks; to develop and implement a process to monitor and measure their safety performance; and to engage in safety promotion through training and communication. An overview of the PTASP is available here: https://www.transit.dot.gov/PTASP.

This new PTASP rule provides a flexible approach to evaluating the safety impacts of automated buses. **FTA recognizes that operating domains and vehicle types and capabilities differ significantly. That is why FTA is not proposing a one-size-fits-all approach**

[22] 23 CFR 655.603

[23] 49 U.S.C. § 5329

[24] Federal Transit Administration, National Public Transportation Safety Plan (Washington: Federal Transit Administration, 2007), www.transit.dot.gov/regulations-and-guidance/safety/national-public-transportation-safety-plan.

[25] 49 C.F.R. Part 673

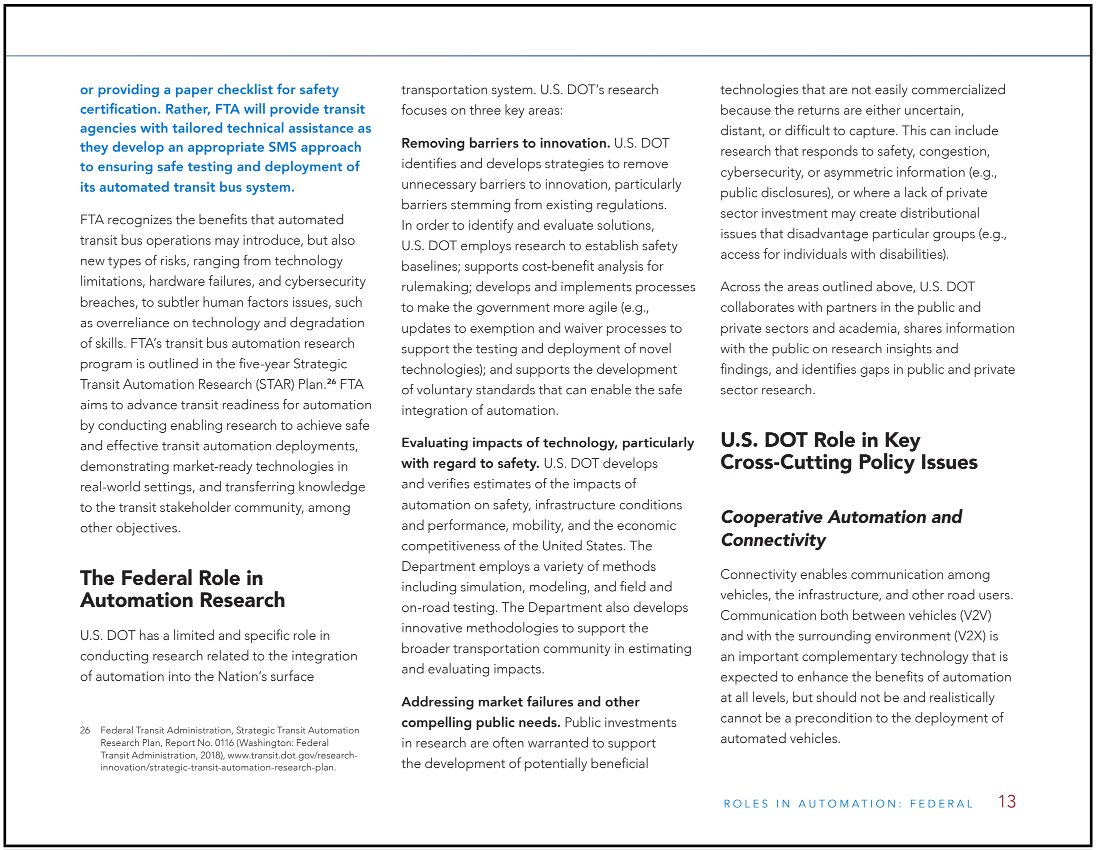
Disability, Accessibility, and Universal Design

Automation presents enormous potential for improving the mobility of travelers with disabilities. Through the Accessible Transportation Technologies Research Initiative (ATTRI), the Department is initiating efforts to partner with the U.S. Department of Labor (DOL), U.S. Department of Health and Human Services (HHS), and the broader disability community to focus research efforts and initiatives on areas where market incentives may otherwise lead to underinvestment.

ATTRI focuses on emerging research, prototyping, and integrated demonstrations with the goal of enabling people to travel independently and conveniently, regardless of their individual abilities. ATTRI research focuses on removing barriers to transportation for people with disabilities, veterans with disabilities, and older adults, with particular attention to those with mobility, cognitive, vision, and hearing disabilities. By leveraging principles of universal design and inclusive information and communication technology, these efforts are targeting solutions that could be transformative for independent mobility.

ATTRI applications in development include wayfinding and navigation, pre-trip concierge and virtualization, safe intersection crossing, and robotics and automation. Automated vehicles and robotics are expected to improve mobility for those unable or unwilling to drive and enhance independent and spontaneous travel capabilities for travelers with disabilities. One area of particular interest among public transit agencies is exploring the use of vehicle automation to solve first mile/last mile mobility issues, possibly providing connections for all travelers to existing public transportation or other transportation hubs.

In addition, machine vision, artificial intelligence (AI), assistive robots, and facial recognition software solving a variety of travel-related issues for persons with disabilities in vehicles, devices, and terminals, are also included to create virtual caregivers/concierge services and other such applications to guide travelers and assist in decision making.



or providing a paper checklist for safety certification. Rather, FTA will provide transit agencies with tailored technical assistance as they develop an appropriate SMS approach to ensuring safe testing and deployment of its automated transit bus system.

FTA recognizes the benefits that automated transit bus operations may introduce, but also new types of risks, ranging from technology limitations, hardware failures, and cybersecurity breaches, to subtler human factors issues, such as overreliance on technology and degradation of skills. FTA's transit bus automation research program is outlined in the five-year Strategic Transit Automation Research (STAR) Plan.[26] FTA aims to advance transit readiness for automation by conducting enabling research to achieve safe and effective transit automation deployments, demonstrating market-ready technologies in real-world settings, and transferring knowledge to the transit stakeholder community, among other objectives.

The Federal Role in Automation Research

U.S. DOT has a limited and specific role in conducting research related to the integration of automation into the Nation's surface transportation system. U.S. DOT's research focuses on three key areas:

Removing barriers to innovation. U.S. DOT identifies and develops strategies to remove unnecessary barriers to innovation, particularly barriers stemming from existing regulations. In order to identify and evaluate solutions, U.S. DOT employs research to establish safety baselines; supports cost-benefit analysis for rulemaking; develops and implements processes to make the government more agile (e.g., updates to exemption and waiver processes to support the testing and deployment of novel technologies); and supports the development of voluntary standards that can enable the safe integration of automation.

Evaluating impacts of technology, particularly with regard to safety. U.S. DOT develops and verifies estimates of the impacts of automation on safety, infrastructure conditions and performance, mobility, and the economic competitiveness of the United States. The Department employs a variety of methods including simulation, modeling, and field and on-road testing. The Department also develops innovative methodologies to support the broader transportation community in estimating and evaluating impacts.

Addressing market failures and other compelling public needs. Public investments in research are often warranted to support the development of potentially beneficial technologies that are not easily commercialized because the returns are either uncertain, distant, or difficult to capture. This can include research that responds to safety, congestion, cybersecurity, or asymmetric information (e.g., public disclosures), or where a lack of private sector investment may create distributional issues that disadvantage particular groups (e.g., access for individuals with disabilities).

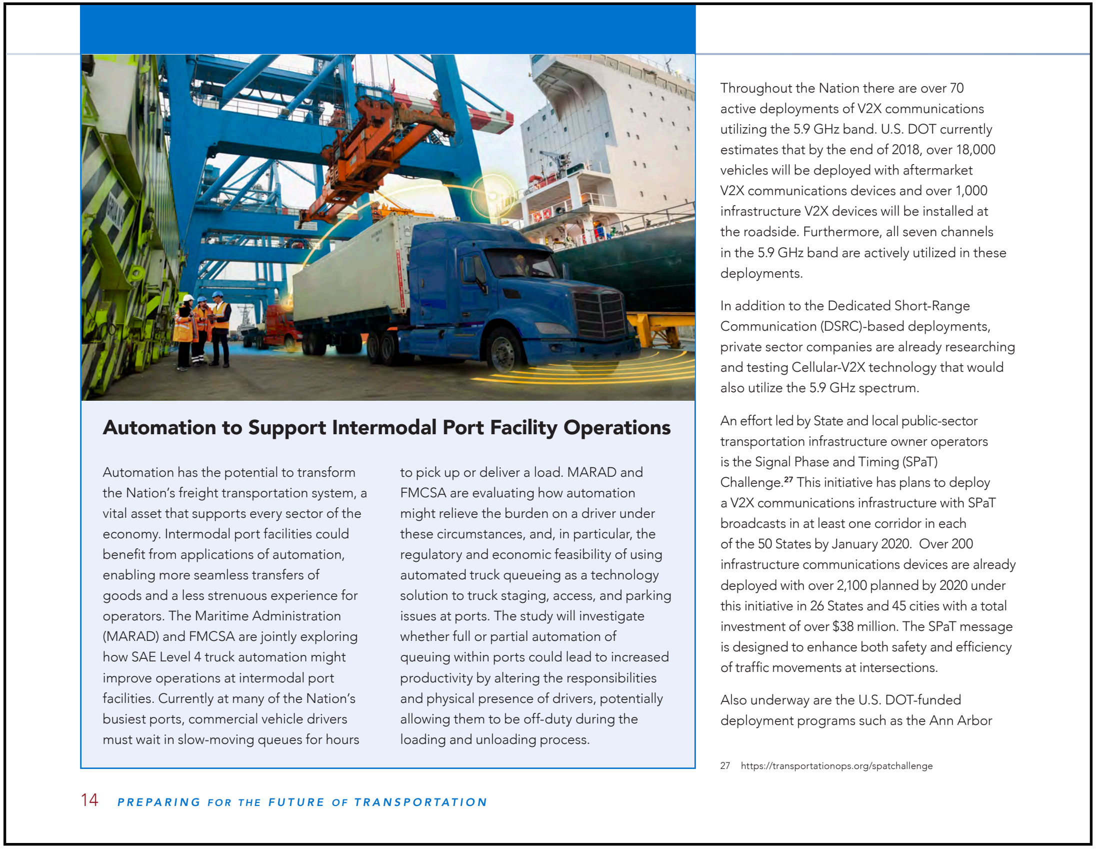
Across the areas outlined above, U.S. DOT collaborates with partners in the public and private sectors and academia, shares information with the public on research insights and findings, and identifies gaps in public and private sector research.

U.S. DOT Role in Key Cross-Cutting Policy Issues

Cooperative Automation and Connectivity

Connectivity enables communication among vehicles, the infrastructure, and other road users. Communication both between vehicles (V2V) and with the surrounding environment (V2X) is an important complementary technology that is expected to enhance the benefits of automation at all levels, but should not be and realistically cannot be a precondition to the deployment of automated vehicles.

[26] Federal Transit Administration, Strategic Transit Automation Research Plan, Report No. 0116 (Washington: Federal Transit Administration, 2018), www.transit.dot.gov/research-innovation/strategic-transit-automation-research-plan.



Throughout the Nation there are over 70 active deployments of V2X communications utilizing the 5.9 GHz band. U.S. DOT currently estimates that by the end of 2018, over 18,000 vehicles will be deployed with aftermarket V2X communications devices and over 1,000 infrastructure V2X devices will be installed at the roadside. Furthermore, all seven channels in the 5.9 GHz band are actively utilized in these deployments.

In addition to the Dedicated Short-Range Communication (DSRC)-based deployments, private sector companies are already researching and testing Cellular-V2X technology that would also utilize the 5.9 GHz spectrum.

An effort led by State and local public-sector transportation infrastructure owner operators is the Signal Phase and Timing (SPaT) Challenge.[27] This initiative has plans to deploy a V2X communications infrastructure with SPaT broadcasts in at least one corridor in each of the 50 States by January 2020. Over 200 infrastructure communications devices are already deployed with over 2,100 planned by 2020 under this initiative in 26 States and 45 cities with a total investment of over $38 million. The SPaT message is designed to enhance both safety and efficiency of traffic movements at intersections.

Also underway are the U.S. DOT-funded deployment programs such as the Ann Arbor

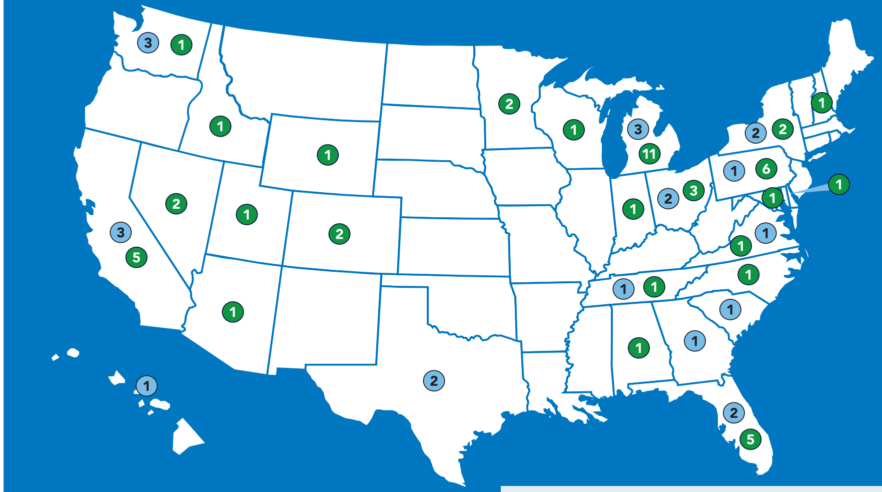
Automation to Support Intermodal Port Facility Operations

Automation has the potential to transform the Nation's freight transportation system, a vital asset that supports every sector of the economy. Intermodal port facilities could benefit from applications of automation, enabling more seamless transfers of goods and a less strenuous experience for operators. The Maritime Administration (MARAD) and FMCSA are jointly exploring how SAE Level 4 truck automation might improve operations at intermodal port facilities. Currently at many of the Nation's busiest ports, commercial vehicle drivers must wait in slow-moving queues for hours to pick up or deliver a load. MARAD and FMCSA are evaluating how automation might relieve the burden on a driver under these circumstances, and, in particular, the regulatory and economic feasibility of using automated truck queueing as a technology solution to truck staging, access, and parking issues at ports. The study will investigate whether full or partial automation of queuing within ports could lead to increased productivity by altering the responsibilities and physical presence of drivers, potentially allowing them to be off-duty during the loading and unloading process.

[27] https://transportationops.org/spatchallenge

Planned and Operational Connected Vehicle Deployments

Where Infrastructure and In-Vehicle Units are Planned or In Use



○ Planned Projects

● Operational Projects

Source: USDOT September 2018

	Infrastructure Units	In-Vehicle Units
Operational (52 Projects)*	2,044	3,340
Planned (23 projects)*, **	242	0
Total	2,286	3,340

* Projects shown include those sponsored by U.S. DOT and others.
** Device numbers for many of the planned projects are currently unavailable.

Cooperative Automation

FHWA is conducting research to measure the efficiency and safety benefits of augmenting automated vehicle capabilities with connected vehicle technologies to enable cooperative automation. Cooperative automation allows automated vehicles to communicate with other vehicles and the infrastructure to coordinate movements and increase efficiency and safety. It uses a range of automation capabilities, including automation technologies at SAE Level 1 and Level 2. Examples of cooperative automation applications include:

- Vehicle platooning to enable safe close following between vehicles and improve highway capacity.
- Speed harmonization using wireless speed control to reduce bottleneck conditions.
- Cooperative lane change and merge functions to mitigate traffic disruptions at interchanges.
- Coordination of signalized intersection approach and departure, using Signal Phase and Timing (SPaT) data to enable automated vehicles to enter and exit signalized intersections safely and efficiently, to mitigate delays and reduce fuel consumption.

Current activities focus on technical assessments, traffic modeling, and proof-of-concept/prototype tests to understand how to improve safety, smooth traffic flow, and reduce fuel consumption. FHWA is partnering with automotive manufacturers to further develop these concepts and is conducting modeling and analysis of corridors in several States. FHWA may pursue further proof-of-concept testing on test tracks and on public roads in the future. Additionally, studies are underway to consider how early automation applications like lane keeping and adaptive cruise control are being used and accepted by everyday drivers.

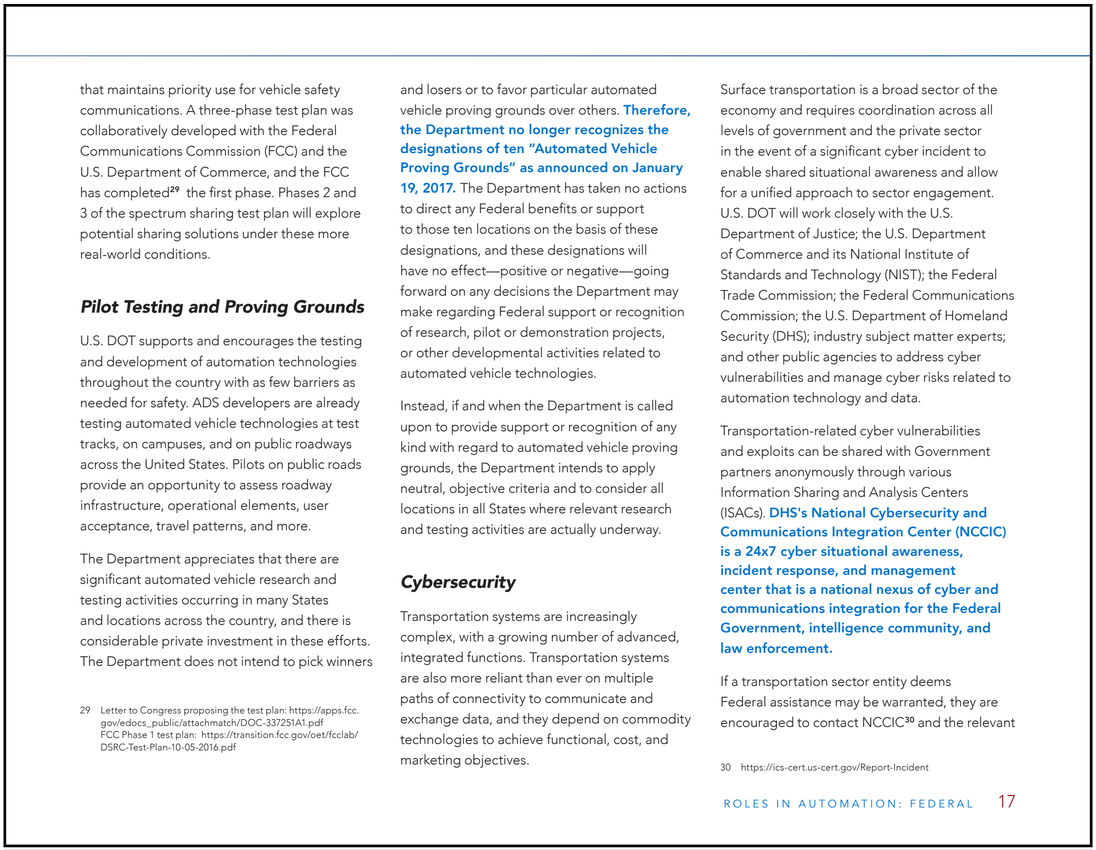
Connected Vehicle Environment, Connected Vehicle Pilots Program, and the Advanced Transportation and Congestion Management Technologies Deployment Program, which have combined over $150 million in Federal and State funding to deploy V2X communications. Finally, states such as Colorado are combining Federal-aid highway program funding with State funding ($72 million) to deploy V2X communications throughout the State highways by 2021.[28]

Over the past 20 years, the U.S. DOT has invested over $700 million in research and development of V2X through partnerships with industry and state/local governments. As a result of these investments and partnerships, V2X technology is on the verge of wide-scale deployment across the Nation.

The Department encourages the automotive industry, wireless technology companies, IOOs, and other stakeholders to continue developing technologies that leverage the 5.9 GHz spectrum for transportation safety benefits. Yet, the Department does not promote any particular technology over another. The Department also encourages the development of connected infrastructure because such technologies offer the potential to improve safety and efficiency. As IOOs consider enabling V2X deployment in their region, the Department encourages IOOs to engage with the U.S. DOT for guidance and assistance.

As part of this approach, U.S. DOT is continuing its work to preserve the ability for transportation safety applications to function in the 5.9 GHz spectrum while exploring methods for sharing the spectrum with other users in a manner

[28] https://www.codot.gov/about/transportation-commission/documents/2018-agendas-and-supporting-documents/june-2018/7-tech-committee.pdf



that maintains priority use for vehicle safety communications. A three-phase test plan was collaboratively developed with the Federal Communications Commission (FCC) and the U.S. Department of Commerce, and the FCC has completed[29] the first phase. Phases 2 and 3 of the spectrum sharing test plan will explore potential sharing solutions under these more real-world conditions.

Pilot Testing and Proving Grounds

U.S. DOT supports and encourages the testing and development of automation technologies throughout the country with as few barriers as needed for safety. ADS developers are already testing automated vehicle technologies at test tracks, on campuses, and on public roadways across the United States. Pilots on public roads provide an opportunity to assess roadway infrastructure, operational elements, user acceptance, travel patterns, and more.

The Department appreciates that there are significant automated vehicle research and testing activities occurring in many States and locations across the country, and there is considerable private investment in these efforts. The Department does not intend to pick winners and losers or to favor particular automated vehicle proving grounds over others. **Therefore, the Department no longer recognizes the designations of ten "Automated Vehicle Proving Grounds" as announced on January 19, 2017.** The Department has taken no actions to direct any Federal benefits or support to those ten locations on the basis of these designations, and these designations will have no effect—positive or negative—going forward on any decisions the Department may make regarding Federal support or recognition of research, pilot or demonstration projects, or other developmental activities related to automated vehicle technologies.

Instead, if and when the Department is called upon to provide support or recognition of any kind with regard to automated vehicle proving grounds, the Department intends to apply neutral, objective criteria and to consider all locations in all States where relevant research and testing activities are actually underway.

Cybersecurity

Transportation systems are increasingly complex, with a growing number of advanced, integrated functions. Transportation systems are also more reliant than ever on multiple paths of connectivity to communicate and exchange data, and they depend on commodity technologies to achieve functional, cost, and marketing objectives.

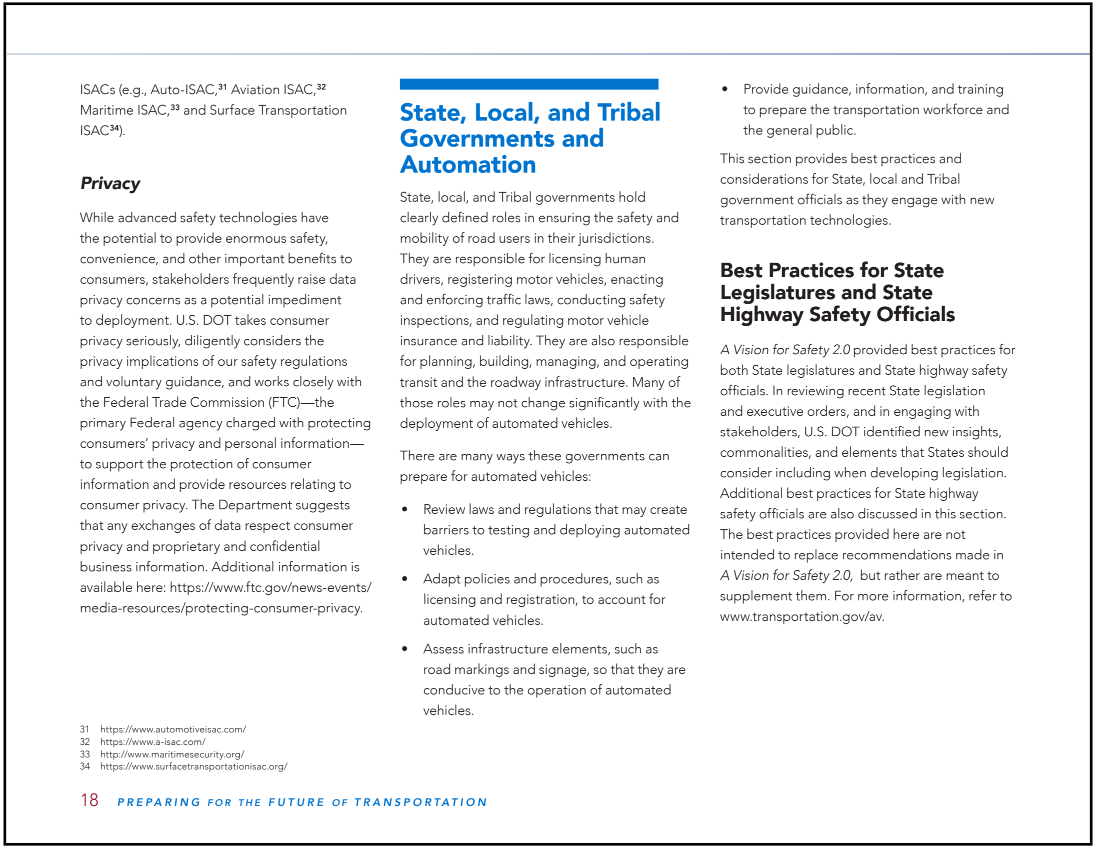
Surface transportation is a broad sector of the economy and requires coordination across all levels of government and the private sector in the event of a significant cyber incident to enable shared situational awareness and allow for a unified approach to sector engagement. U.S. DOT will work closely with the U.S. Department of Justice; the U.S. Department of Commerce and its National Institute of Standards and Technology (NIST); the Federal Trade Commission; the Federal Communications Commission; the U.S. Department of Homeland Security (DHS); industry subject matter experts; and other public agencies to address cyber vulnerabilities and manage cyber risks related to automation technology and data.

Transportation-related cyber vulnerabilities and exploits can be shared with Government partners anonymously through various Information Sharing and Analysis Centers (ISACs). **DHS's National Cybersecurity and Communications Integration Center (NCCIC) is a 24x7 cyber situational awareness, incident response, and management center that is a national nexus of cyber and communications integration for the Federal Government, intelligence community, and law enforcement.**

If a transportation sector entity deems Federal assistance may be warranted, they are encouraged to contact NCCIC[30] and the relevant

[29] Letter to Congress proposing the test plan: https://apps.fcc.gov/edocs_public/attachmatch/DOC-337251A1.pdf
FCC Phase 1 test plan: https://transition.fcc.gov/oet/fcclab/DSRC-Test-Plan-10-05-2016.pdf

[30] https://ics-cert.us-cert.gov/Report-Incident



ISACs (e.g., Auto-ISAC,[31] Aviation ISAC,[32] Maritime ISAC,[33] and Surface Transportation ISAC[34]).

Privacy

While advanced safety technologies have the potential to provide enormous safety, convenience, and other important benefits to consumers, stakeholders frequently raise data privacy concerns as a potential impediment to deployment. U.S. DOT takes consumer privacy seriously, diligently considers the privacy implications of our safety regulations and voluntary guidance, and works closely with the Federal Trade Commission (FTC)—the primary Federal agency charged with protecting consumers' privacy and personal information—to support the protection of consumer information and provide resources relating to consumer privacy. The Department suggests that any exchanges of data respect consumer privacy and proprietary and confidential business information. Additional information is available here: https://www.ftc.gov/news-events/media-resources/protecting-consumer-privacy.

State, Local, and Tribal Governments and Automation

State, local, and Tribal governments hold clearly defined roles in ensuring the safety and mobility of road users in their jurisdictions. They are responsible for licensing human drivers, registering motor vehicles, enacting and enforcing traffic laws, conducting safety inspections, and regulating motor vehicle insurance and liability. They are also responsible for planning, building, managing, and operating transit and the roadway infrastructure. Many of those roles may not change significantly with the deployment of automated vehicles.

There are many ways these governments can prepare for automated vehicles:

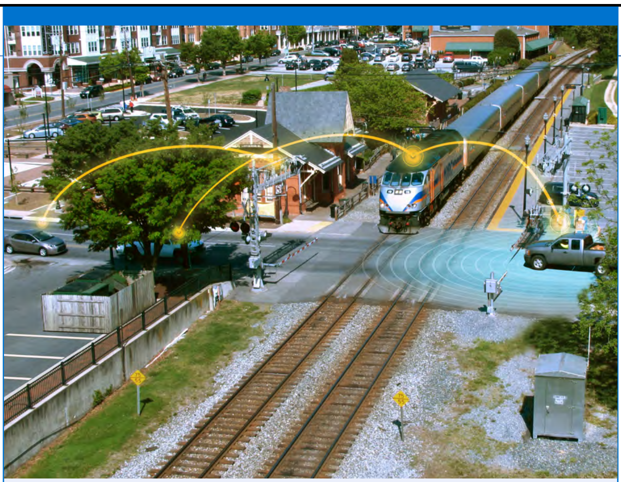
- Review laws and regulations that may create barriers to testing and deploying automated vehicles.
- Adapt policies and procedures, such as licensing and registration, to account for automated vehicles.
- Assess infrastructure elements, such as road markings and signage, so that they are conducive to the operation of automated vehicles.
- Provide guidance, information, and training to prepare the transportation workforce and the general public.

This section provides best practices and considerations for State, local and Tribal government officials as they engage with new transportation technologies.

Best Practices for State Legislatures and State Highway Safety Officials

A Vision for Safety 2.0 provided best practices for both State legislatures and State highway safety officials. In reviewing recent State legislation and executive orders, and in engaging with stakeholders, U.S. DOT identified new insights, commonalities, and elements that States should consider including when developing legislation. Additional best practices for State highway safety officials are also discussed in this section. The best practices provided here are not intended to replace recommendations made in *A Vision for Safety 2.0,* but rather are meant to supplement them. For more information, refer to www.transportation.gov/av.

[31] https://www.automotiveisac.com/
[32] https://www.a-isac.com/
[33] http://www.maritimesecurity.org/
[34] https://www.surfacetransportationisac.org/



Automated Vehicles at Rail Crossings

To explore the interaction between automated vehicles and highway-rail grade crossings and identify what information automated vehicles will need in order to negotiate highway-rail intersections, the Federal Railroad Administration (FRA) has conducted a literature review, engaged with stakeholders, and used scenarios to develop and demonstrate a concept of operations, including system requirements (technology and sensors).

A broad stakeholder set was identified to represent researchers, manufacturers, transit agencies, and infrastructure owner-operators, among others. Currently, FRA is expanding the research with U.S. DOT partners and the Association of American Railroads to develop a closed loop safety system to support the safe interaction of connected and automated vehicles with grade crossings.

Best Practices for State Legislatures

States are taking differing legislative approaches and have enacted varying laws related to testing and operating automated vehicles. U.S. DOT regularly monitors legislative activities in order to support the development of a consistent national framework for automated vehicle legislation.

A Vision for Safety 2.0 recommended that State legislators follow best practices, such as providing a technology-neutral environment, licensing and registration procedures, and reporting and communications methods for public safety officials. States should consider reviewing and potentially modifying traffic laws and regulations that may be barriers to automated vehicles. For example, several States have following distance laws that prohibit trucks from following too closely to each other, effectively prohibiting automated truck platooning applications.

In addition to the best practices identified in *A Vision for Safety 2.0*, the Department recommends that State officials consider the following safety-related best practices when crafting automated vehicle legislation:

Engage U.S. DOT on legislative technical assistance. State legislatures are encouraged to routinely engage U.S. DOT on legislative activities related to multimodal automation

safety. State legislatures may want to first determine if there is a need for State legislation. Unnecessary or overly prescriptive **State requirements could create unintended barriers for the testing, deployment, and operations of advanced vehicle safety technologies.** U.S. DOT stands ready to provide technical assistance to States on request.

Adopt terminology defined through voluntary technical standards. Different use and interpretations of terminology regarding automated vehicles can be confusing for the public, State and local agencies, and industry. In the interest of supporting consistent terminology, State legislatures may want to use terminology already being developed through voluntary, consensus-based, technical standards. SAE terminology on automation represents one example and includes terms such as ADS, the Dynamic Driving Task (DDT), minimal risk conditions, and ODD.

Assess State roadway readiness. States may want to assess roadway readiness for automated vehicles, as such assessments could help infrastructure for automated vehicles, while improving safety for drivers today. Automated vehicle developers are designing their technologies with the assumption that these technologies will need to function with existing infrastructure. There is general agreement that greater uniformity and quality of road markings, signage, and pavement condition would be beneficial for both human drivers and automated vehicles.

Best Practices for State Highway Safety Officials

States are responsible for reducing traffic crashes and resulting deaths, injuries, and property damage for all road users in their jurisdictions. States use this authority to establish and maintain highway safety programs addressing driver education and testing, licensing, pedestrian safety, and vehicle registration and inspection. States also use this authority to address traffic control, highway design and maintenance, crash prevention, investigation and recordkeeping, and law enforcement and emergency service considerations.

The following best practices build on those identified in *A Vision for Safety 2.0* and provide a framework for States looking for assistance in developing procedures and conditions for the operation of automated vehicles on public roadways. For additional best practices, see Section 2 of *A Vision for Safety 2.0.*

Consider test driver training and licensing procedures for test vehicles. States may consider minimum requirements for test drivers who operate test vehicles at different automation levels. States may want to coordinate and collaborate with a broad and diverse set of stakeholders when developing and defining jurisdictional guidelines for safe testing and deployment of automated vehicles.

Recognize issues unique to entities offering automated mobility as a service. Automated mobility providers are exploring models to move people and goods using automated vehicle technology. States may consider identifying and addressing issues that are unique to companies providing mobility as a service using automated vehicle technologies. These could include such issues as congestion or the transportation of minors, persons with disabilities, and older individuals.

Considerations for Infrastructure Owners and Operators

Infrastructure owners and operators are involved in the planning, design, construction, maintenance, and operation of the roadway infrastructure. Infrastructure owners and operators have expressed interest in more information and guidance on how to prepare for automated vehicle deployment and testing on public roadways. FHWA is conducting the National Dialogue on Highway Automation, a series of workshops with partners, stakeholders, and the public to obtain input regarding the safe

and efficient integration of automated vehicles into the roadway system.[35] U.S. DOT provides the following considerations for infrastructure owners and operators, including State DOTs, metropolitan planning organizations (MPOs), and local agencies. FHWA, in particular, will continue to update these considerations as informed by continued research efforts, stakeholder engagement, and testing. Suggested considerations include:

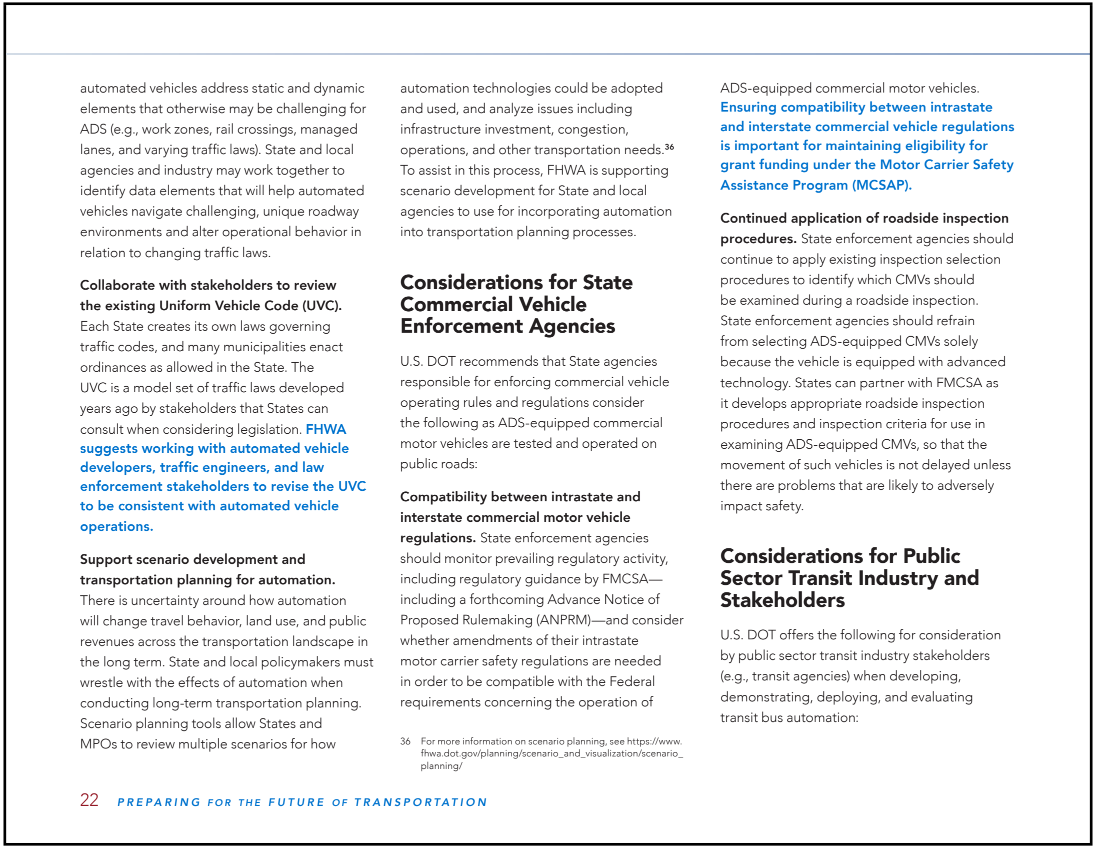
Support safe testing and operations of automated vehicles on public roadways. State DOTs and local agencies want to understand under what conditions automated vehicles can safely operate in automated mode and how they will affect the highway infrastructure and surrounding communities. Where testing is taking place, State and local agencies should consider ways to establish consistent cross-jurisdictional approaches and work with first responders to develop commonly understood traffic law enforcement practices and emergency response plans for automated vehicle testing and operation.

Learn from testing and pilots to support highway system readiness. State and local agencies may consider collaborating with automated vehicle developers and testers to identify potential infrastructure requirements that support readiness for automated vehicles and to understand their expectations for automated vehicle operations under varying roadway and operational conditions. This interaction could assist with identifying what balance of capabilities (for both vehicles and the roadway) promotes safe and efficient operations of automated vehicles. Testing, research, and pilot programs can help State and local agencies understand automation and identify opportunities to inform transportation planning, infrastructure design, and traffic operations management.

Build organizational capacity to prepare for automated vehicles in communities. State and local agencies may need to assess their workforce capacity and training needs to address new issues that emerge from having automated vehicles on public roads. State and local agencies will want to work with peers, industry, associations, the research community, and FHWA to build knowledge of automated vehicle technologies and identify technical assistance resources.

Identify data needs and opportunities to exchange data. The exchange of data and information in the roadway environment can help

[35] More information can be found at https://ops.fhwa.dot.gov/automationdialogue/



automated vehicles address static and dynamic elements that otherwise may be challenging for ADS (e.g., work zones, rail crossings, managed lanes, and varying traffic laws). State and local agencies and industry may work together to identify data elements that will help automated vehicles navigate challenging, unique roadway environments and alter operational behavior in relation to changing traffic laws.

Collaborate with stakeholders to review the existing Uniform Vehicle Code (UVC). Each State creates its own laws governing traffic codes, and many municipalities enact ordinances as allowed in the State. The UVC is a model set of traffic laws developed years ago by stakeholders that States can consult when considering legislation. **FHWA suggests working with automated vehicle developers, traffic engineers, and law enforcement stakeholders to revise the UVC to be consistent with automated vehicle operations.**

Support scenario development and transportation planning for automation. There is uncertainty around how automation will change travel behavior, land use, and public revenues across the transportation landscape in the long term. State and local policymakers must wrestle with the effects of automation when conducting long-term transportation planning. Scenario planning tools allow States and MPOs to review multiple scenarios for how automation technologies could be adopted and used, and analyze issues including infrastructure investment, congestion, operations, and other transportation needs.[36] To assist in this process, FHWA is supporting scenario development for State and local agencies to use for incorporating automation into transportation planning processes.

Considerations for State Commercial Vehicle Enforcement Agencies

U.S. DOT recommends that State agencies responsible for enforcing commercial vehicle operating rules and regulations consider the following as ADS-equipped commercial motor vehicles are tested and operated on public roads:

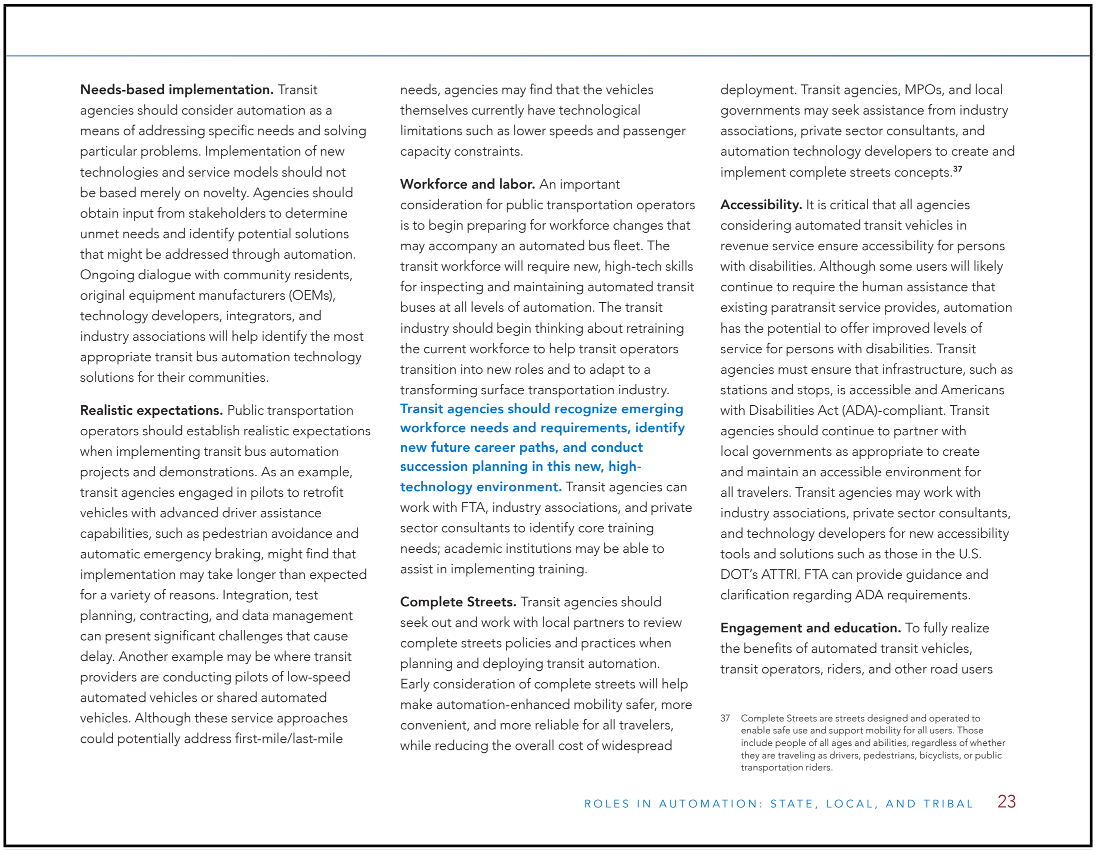
Compatibility between intrastate and interstate commercial motor vehicle regulations. State enforcement agencies should monitor prevailing regulatory activity, including regulatory guidance by FMCSA—including a forthcoming Advance Notice of Proposed Rulemaking (ANPRM)—and consider whether amendments of their intrastate motor carrier safety regulations are needed in order to be compatible with the Federal requirements concerning the operation of ADS-equipped commercial motor vehicles. **Ensuring compatibility between intrastate and interstate commercial vehicle regulations is important for maintaining eligibility for grant funding under the Motor Carrier Safety Assistance Program (MCSAP).**

Continued application of roadside inspection procedures. State enforcement agencies should continue to apply existing inspection selection procedures to identify which CMVs should be examined during a roadside inspection. State enforcement agencies should refrain from selecting ADS-equipped CMVs solely because the vehicle is equipped with advanced technology. States can partner with FMCSA as it develops appropriate roadside inspection procedures and inspection criteria for use in examining ADS-equipped CMVs, so that the movement of such vehicles is not delayed unless there are problems that are likely to adversely impact safety.

Considerations for Public Sector Transit Industry and Stakeholders

U.S. DOT offers the following for consideration by public sector transit industry stakeholders (e.g., transit agencies) when developing, demonstrating, deploying, and evaluating transit bus automation:

[36] For more information on scenario planning, see https://www.fhwa.dot.gov/planning/scenario_and_visualization/scenario_planning/



Needs-based implementation. Transit agencies should consider automation as a means of addressing specific needs and solving particular problems. Implementation of new technologies and service models should not be based merely on novelty. Agencies should obtain input from stakeholders to determine unmet needs and identify potential solutions that might be addressed through automation. Ongoing dialogue with community residents, original equipment manufacturers (OEMs), technology developers, integrators, and industry associations will help identify the most appropriate transit bus automation technology solutions for their communities.

Realistic expectations. Public transportation operators should establish realistic expectations when implementing transit bus automation projects and demonstrations. As an example, transit agencies engaged in pilots to retrofit vehicles with advanced driver assistance capabilities, such as pedestrian avoidance and automatic emergency braking, might find that implementation may take longer than expected for a variety of reasons. Integration, test planning, contracting, and data management can present significant challenges that cause delay. Another example may be where transit providers are conducting pilots of low-speed automated vehicles or shared automated vehicles. Although these service approaches could potentially address first-mile/last-mile needs, agencies may find that the vehicles themselves currently have technological limitations such as lower speeds and passenger capacity constraints.

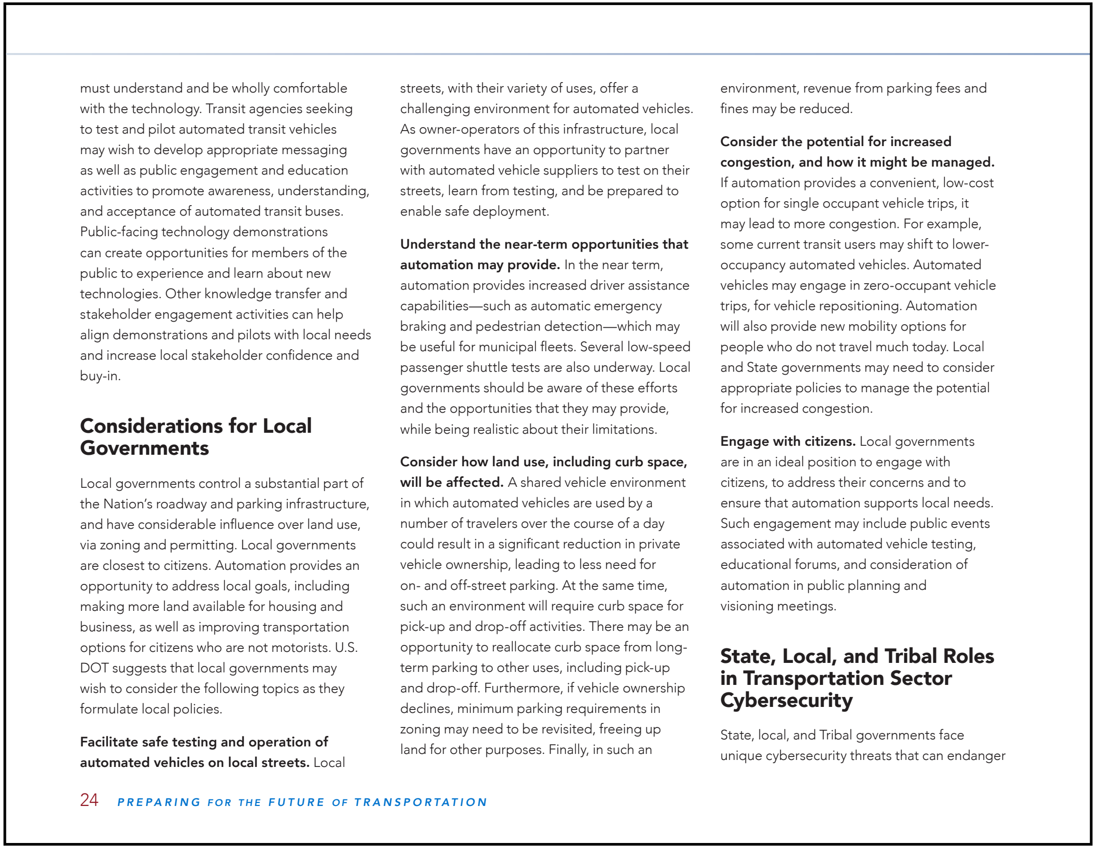
Workforce and labor. An important consideration for public transportation operators is to begin preparing for workforce changes that may accompany an automated bus fleet. The transit workforce will require new, high-tech skills for inspecting and maintaining automated transit buses at all levels of automation. The transit industry should begin thinking about retraining the current workforce to help transit operators transition into new roles and to adapt to a transforming surface transportation industry. **Transit agencies should recognize emerging workforce needs and requirements, identify new future career paths, and conduct succession planning in this new, high-technology environment.** Transit agencies can work with FTA, industry associations, and private sector consultants to identify core training needs; academic institutions may be able to assist in implementing training.

Complete Streets. Transit agencies should seek out and work with local partners to review complete streets policies and practices when planning and deploying transit automation. Early consideration of complete streets will help make automation-enhanced mobility safer, more convenient, and more reliable for all travelers, while reducing the overall cost of widespread deployment. Transit agencies, MPOs, and local governments may seek assistance from industry associations, private sector consultants, and automation technology developers to create and implement complete streets concepts.[37]

Accessibility. It is critical that all agencies considering automated transit vehicles in revenue service ensure accessibility for persons with disabilities. Although some users will likely continue to require the human assistance that existing paratransit service provides, automation has the potential to offer improved levels of service for persons with disabilities. Transit agencies must ensure that infrastructure, such as stations and stops, is accessible and Americans with Disabilities Act (ADA)-compliant. Transit agencies should continue to partner with local governments as appropriate to create and maintain an accessible environment for all travelers. Transit agencies may work with industry associations, private sector consultants, and technology developers for new accessibility tools and solutions such as those in the U.S. DOT's ATTRI. FTA can provide guidance and clarification regarding ADA requirements.

Engagement and education. To fully realize the benefits of automated transit vehicles, transit operators, riders, and other road users must understand and be wholly comfortable with the technology. Transit agencies seeking to test and pilot automated transit vehicles may wish to develop appropriate messaging as well as public engagement and education activities to promote awareness, understanding, and acceptance of automated transit buses. Public-facing technology demonstrations can create opportunities for members of the public to experience and learn about new technologies. Other knowledge transfer and stakeholder engagement activities can help align demonstrations and pilots with local needs and increase local stakeholder confidence and buy-in.

[37] Complete Streets are streets designed and operated to enable safe use and support mobility for all users. Those include people of all ages and abilities, regardless of whether they are traveling as drivers, pedestrians, bicyclists, or public transportation riders.



Considerations for Local Governments

Local governments control a substantial part of the Nation's roadway and parking infrastructure, and have considerable influence over land use, via zoning and permitting. Local governments are closest to citizens. Automation provides an opportunity to address local goals, including making more land available for housing and business, as well as improving transportation options for citizens who are not motorists. U.S. DOT suggests that local governments may wish to consider the following topics as they formulate local policies.

Facilitate safe testing and operation of automated vehicles on local streets. Local streets, with their variety of uses, offer a challenging environment for automated vehicles. As owner-operators of this infrastructure, local governments have an opportunity to partner with automated vehicle suppliers to test on their streets, learn from testing, and be prepared to enable safe deployment.

Understand the near-term opportunities that automation may provide. In the near term, automation provides increased driver assistance capabilities—such as automatic emergency braking and pedestrian detection—which may be useful for municipal fleets. Several low-speed passenger shuttle tests are also underway. Local governments should be aware of these efforts and the opportunities that they may provide, while being realistic about their limitations.

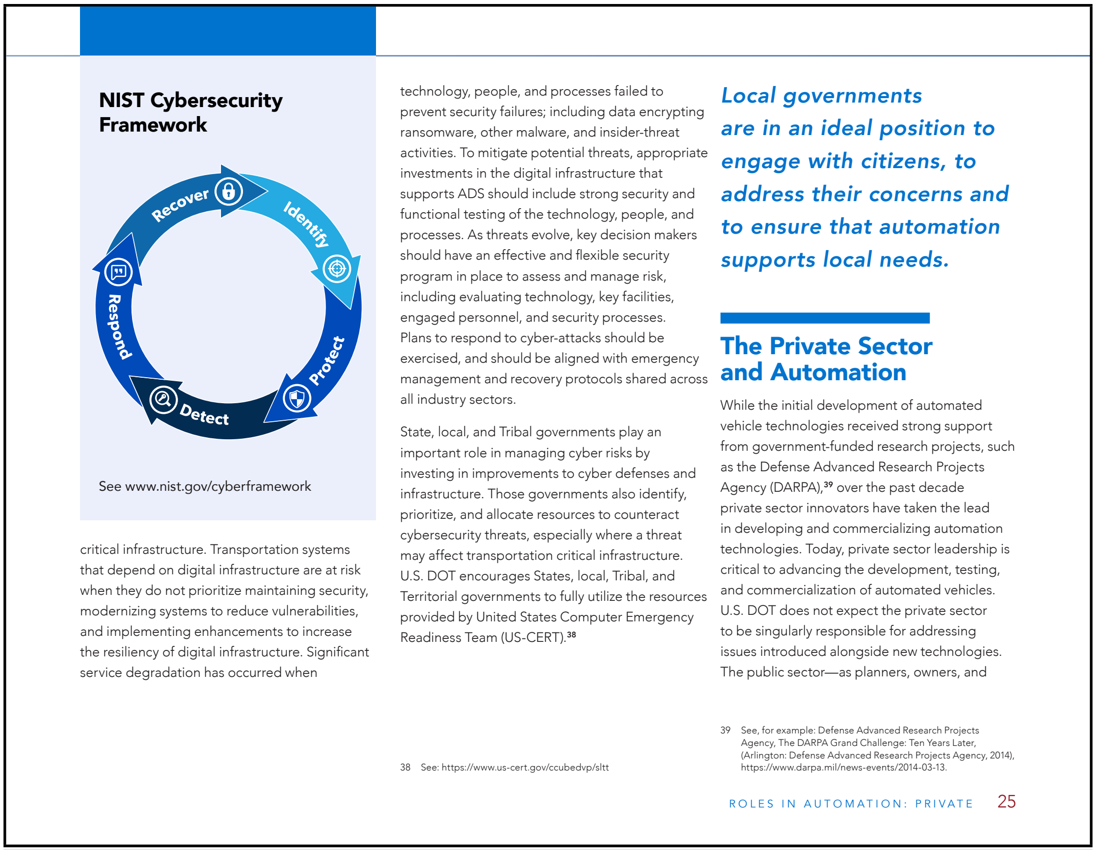
Consider how land use, including curb space, will be affected. A shared vehicle environment in which automated vehicles are used by a number of travelers over the course of a day could result in a significant reduction in private vehicle ownership, leading to less need for on- and off-street parking. At the same time, such an environment will require curb space for pick-up and drop-off activities. There may be an opportunity to reallocate curb space from long-term parking to other uses, including pick-up and drop-off. Furthermore, if vehicle ownership declines, minimum parking requirements in zoning may need to be revisited, freeing up land for other purposes. Finally, in such an environment, revenue from parking fees and fines may be reduced.

Consider the potential for increased congestion, and how it might be managed. If automation provides a convenient, low-cost option for single occupant vehicle trips, it may lead to more congestion. For example, some current transit users may shift to lower-occupancy automated vehicles. Automated vehicles may engage in zero-occupant vehicle trips, for vehicle repositioning. Automation will also provide new mobility options for people who do not travel much today. Local and State governments may need to consider appropriate policies to manage the potential for increased congestion.

Engage with citizens. Local governments are in an ideal position to engage with citizens, to address their concerns and to ensure that automation supports local needs. Such engagement may include public events associated with automated vehicle testing, educational forums, and consideration of automation in public planning and visioning meetings.

State, Local, and Tribal Roles in Transportation Sector Cybersecurity

State, local, and Tribal governments face unique cybersecurity threats that can endanger



NIST Cybersecurity Framework

See www.nist.gov/cyberframework

critical infrastructure. Transportation systems that depend on digital infrastructure are at risk when they do not prioritize maintaining security, modernizing systems to reduce vulnerabilities, and implementing enhancements to increase the resiliency of digital infrastructure. Significant service degradation has occurred when technology, people, and processes failed to prevent security failures; including data encrypting ransomware, other malware, and insider-threat activities. To mitigate potential threats, appropriate investments in the digital infrastructure that supports ADS should include strong security and functional testing of the technology, people, and processes. As threats evolve, key decision makers should have an effective and flexible security program in place to assess and manage risk, including evaluating technology, key facilities, engaged personnel, and security processes. Plans to respond to cyber-attacks should be exercised, and should be aligned with emergency management and recovery protocols shared across all industry sectors.

State, local, and Tribal governments play an important role in managing cyber risks by investing in improvements to cyber defenses and infrastructure. Those governments also identify, prioritize, and allocate resources to counteract cybersecurity threats, especially where a threat may affect transportation critical infrastructure. U.S. DOT encourages States, local, Tribal, and Territorial governments to fully utilize the resources provided by United States Computer Emergency Readiness Team (US-CERT).[38]

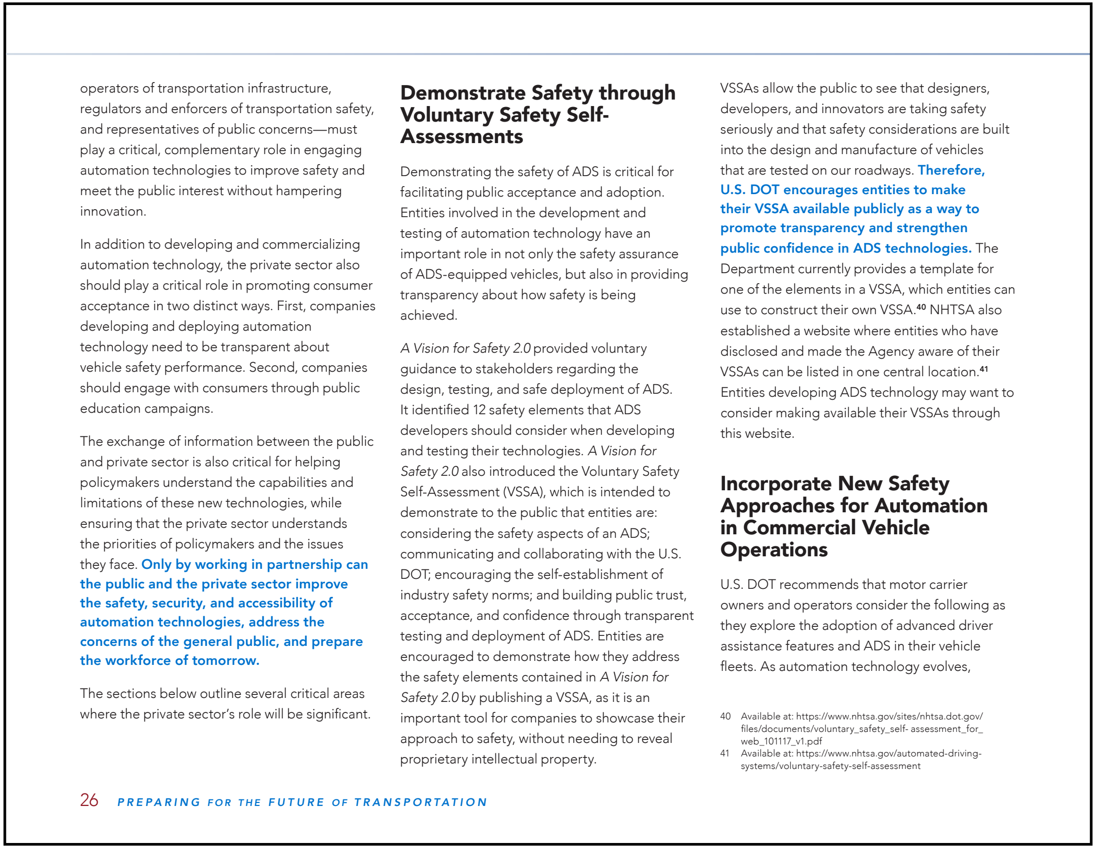
Local governments are in an ideal position to engage with citizens, to address their concerns and to ensure that automation supports local needs.

The Private Sector and Automation

While the initial development of automated vehicle technologies received strong support from government-funded research projects, such as the Defense Advanced Research Projects Agency (DARPA),[39] over the past decade private sector innovators have taken the lead in developing and commercializing automation technologies. Today, private sector leadership is critical to advancing the development, testing, and commercialization of automated vehicles. U.S. DOT does not expect the private sector to be singularly responsible for addressing issues introduced alongside new technologies. The public sector—as planners, owners, and

[38] See: https://www.us-cert.gov/ccubedvp/sltt

[39] See, for example: Defense Advanced Research Projects Agency, The DARPA Grand Challenge: Ten Years Later, (Arlington: Defense Advanced Research Projects Agency, 2014), https://www.darpa.mil/news-events/2014-03-13.



operators of transportation infrastructure, regulators and enforcers of transportation safety, and representatives of public concerns—must play a critical, complementary role in engaging automation technologies to improve safety and meet the public interest without hampering innovation.

In addition to developing and commercializing automation technology, the private sector also should play a critical role in promoting consumer acceptance in two distinct ways. First, companies developing and deploying automation technology need to be transparent about vehicle safety performance. Second, companies should engage with consumers through public education campaigns.

The exchange of information between the public and private sector is also critical for helping policymakers understand the capabilities and limitations of these new technologies, while ensuring that the private sector understands the priorities of policymakers and the issues they face. **Only by working in partnership can the public and the private sector improve the safety, security, and accessibility of automation technologies, address the concerns of the general public, and prepare the workforce of tomorrow.**

The sections below outline several critical areas where the private sector's role will be significant.

Demonstrate Safety through Voluntary Safety Self-Assessments

Demonstrating the safety of ADS is critical for facilitating public acceptance and adoption. Entities involved in the development and testing of automation technology have an important role in not only the safety assurance of ADS-equipped vehicles, but also in providing transparency about how safety is being achieved.

A Vision for Safety 2.0 provided voluntary guidance to stakeholders regarding the design, testing, and safe deployment of ADS. It identified 12 safety elements that ADS developers should consider when developing and testing their technologies. *A Vision for Safety 2.0* also introduced the Voluntary Safety Self-Assessment (VSSA), which is intended to demonstrate to the public that entities are: considering the safety aspects of an ADS; communicating and collaborating with the U.S. DOT; encouraging the self-establishment of industry safety norms; and building public trust, acceptance, and confidence through transparent testing and deployment of ADS. Entities are encouraged to demonstrate how they address the safety elements contained in *A Vision for Safety 2.0* by publishing a VSSA, as it is an important tool for companies to showcase their approach to safety, without needing to reveal proprietary intellectual property.

VSSAs allow the public to see that designers, developers, and innovators are taking safety seriously and that safety considerations are built into the design and manufacture of vehicles that are tested on our roadways. **Therefore, U.S. DOT encourages entities to make their VSSA available publicly as a way to promote transparency and strengthen public confidence in ADS technologies.** The Department currently provides a template for one of the elements in a VSSA, which entities can use to construct their own VSSA.[40] NHTSA also established a website where entities who have disclosed and made the Agency aware of their VSSAs can be listed in one central location.[41] Entities developing ADS technology may want to consider making available their VSSAs through this website.

Incorporate New Safety Approaches for Automation in Commercial Vehicle Operations

U.S. DOT recommends that motor carrier owners and operators consider the following as they explore the adoption of advanced driver assistance features and ADS in their vehicle fleets. As automation technology evolves,

[40] Available at: https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/voluntary_safety_self- assessment_for_web_101117_v1.pdf

[41] Available at: https://www.nhtsa.gov/automated-driving-systems/voluntary-safety-self-assessment

Hazardous Materials Documentation

The Pipeline and Hazardous Materials Safety Administration (PHMSA) is exploring alternatives to longstanding requirements for providing paper documentation to accompany hazmat shipments, while ensuring that the information is readily available to transport workers and emergency responders. **This capability may become increasingly important as transporters of hazardous materials explore the use of automation in their operations.** As motor carriers and railroads explore the use of automation to move hazardous materials, the ability to create electronic documentation also raises the potential to electronically transmit information to first responders before they arrive at an incident. PHMSA is also collaborating with the Environmental Protection Agency on the development of an e-manifest system that will digitize the exchange of information on hazardous material shipments.

FMCSA and PHMSA plan to solicit stakeholder input and provide more detailed guidance regarding the use of ADS in commercial vehicle operations.

System knowledge. If a motor carrier of passengers or property plans to begin operating a commercial motor vehicle equipped with driver-assist systems and/or ADS, the motor carrier's personnel should understand the capabilities and limitations of these systems, as well as ODD limitations (e.g., the types of roadway environments or environmental conditions under which they can operate). The motor carrier should also ask the equipment's manufacturer about the capabilities and limitations of these systems. Motor carriers may also wish to inquire about whether the manufacturer has completed a voluntary safety self-assessment, as described in *A Vision for Safety 2.0.*

System functionality. Motor carriers should ensure the driver assist system and/or ADS is functioning properly before activating these systems. This functionality should be able to be validated during a roadside inspection.

System training. Motor carriers should implement a training program to familiarize fleet managers, maintenance personnel, and drivers with the equipment and how it operates, including the procedures to follow in the event of an ADS malfunction.

Equipment maintenance. Motor carriers should be aware of maintenance requirements of driver-assist systems and/or ADS to enable safe and optimal operation. This includes understanding self-diagnostic capabilities of the system and the status or error messages the system may display.

Information exchange. Motor carriers should be aware that under certain situations such as a safety inspection or roadway crash, it may be necessary to exchange critical safety-oriented vehicle performance data with Federal and State officials. The motor carrier should maintain records of the systems it is using, the training provided, and the operation of those vehicles.

Safety inspections. Motor carriers should be prepared to interact and cooperate with roadside and other safety inspections of driver assist systems and ADS. This includes responding to law enforcement instructions, resolving any identified mechanical or software malfunction, implementing the equipment's safe shutdown procedures, and demonstrating system functionality.

Develop Safe and Accessible Transit Buses and Applications: Considerations for Private Sector Transit Industry

U.S. DOT offers the following considerations for private sector transit industry stakeholders when developing, demonstrating, deploying, and evaluating transit bus automation:

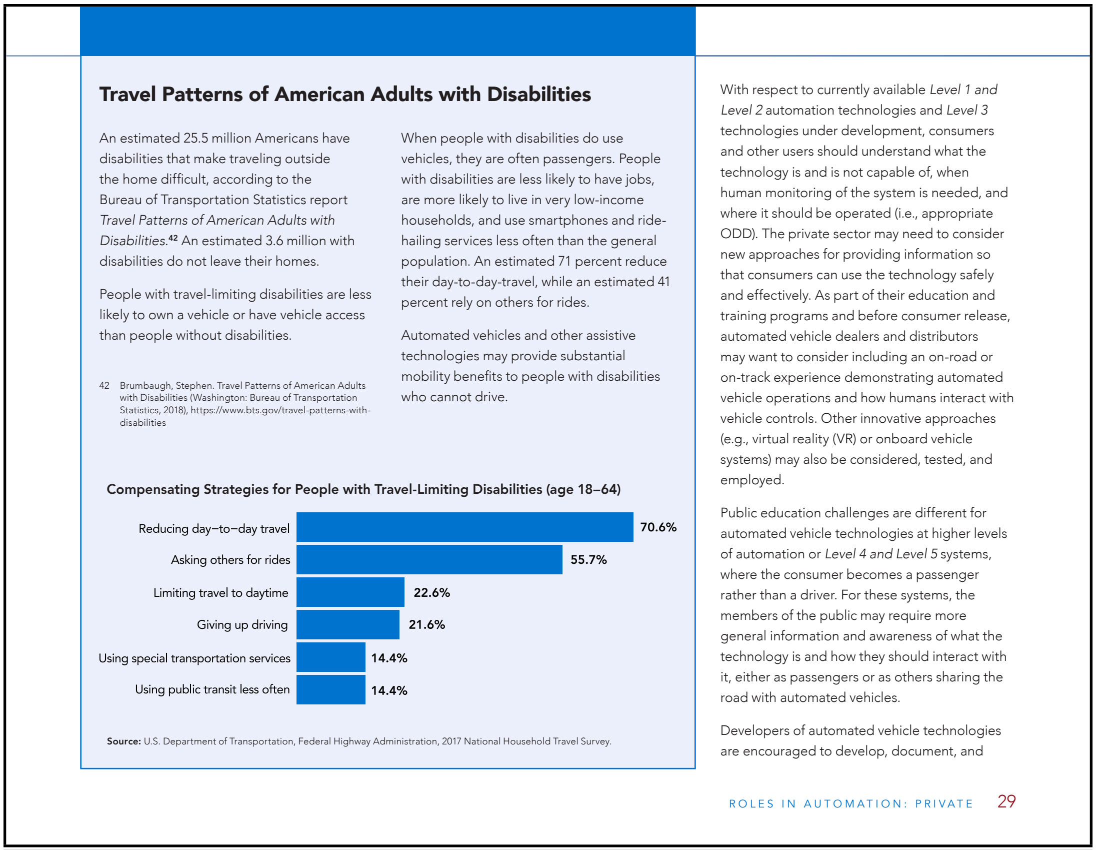
Accessibility. It is important to think about how to make automated vehicles and their technological capabilities accessible to persons with disabilities (including those with physical, sensory, and cognitive impairments) early in the design process. This vital element is more easily integrated at the initial stages of vehicle research and development, rather than trying to incorporate it into the design through retrofits, which may be more difficult. Bus OEMs, technology developers, and integrators should work with transit agencies, industry associations, and the disability community to obtain input on functional and performance needs as well as the consequent human factors considerations. The Federal Government (e.g., FTA) can provide guidance and clarification with respect to the requirements of ADA.

Human factors. Consider human factors in the design of buses and vehicles for all levels of automation—for all participants in the system (transit operators, passengers, and other road users). The interaction between human and machine, ease of use, and comprehensibility of human-machine interfaces (HMI) should be explored thoroughly, particularly with respect to maintaining safety under all operating conditions. Where possible, technology companies should partner with transit agencies and passenger organizations to test various user-interface technologies and designs.

Testing. Open a dialogue and seek a collaborative relationship with FTA when developing and testing new bus technologies and products. FTA can provide guidance, feedback, and clarification on policies, requirements, and recommendations as they pertain to transit automation.

Provide Information to the Public

The understanding of automation technologies varies considerably across the general public, caused in part by a lack of consistency in terminology and confusion about the technology's limitations. The public needs accurate sources of information regarding automation to better understand the technology so that they can use it safely and make informed decisions about its integration. This can be done through direct communications with consumers and other users, demonstrations, public outreach in areas where vehicles are being tested, and a variety of other means.



Travel Patterns of American Adults with Disabilities

An estimated 25.5 million Americans have disabilities that make traveling outside the home difficult, according to the Bureau of Transportation Statistics report *Travel Patterns of American Adults with Disabilities.*[42] An estimated 3.6 million with disabilities do not leave their homes.

People with travel-limiting disabilities are less likely to own a vehicle or have vehicle access than people without disabilities.

When people with disabilities do use vehicles, they are often passengers. People with disabilities are less likely to have jobs, are more likely to live in very low-income households, and use smartphones and ride-hailing services less often than the general population. An estimated 71 percent reduce their day-to-day-travel, while an estimated 41 percent rely on others for rides.

Automated vehicles and other assistive technologies may provide substantial mobility benefits to people with disabilities who cannot drive.

[42] Brumbaugh, Stephen. Travel Patterns of American Adults with Disabilities (Washington: Bureau of Transportation Statistics, 2018), https://www.bts.gov/travel-patterns-with-disabilities

Compensating Strategies for People with Travel-Limiting Disabilities (age 18–64)

Reducing day–to–day travel 70.6%

Asking others for rides 55.7%

Limiting travel to daytime 22.6%

Giving up driving 21.6%

Using special transportation services 14.4%

Using public transit less often 14.4%

Source: U.S. Department of Transportation, Federal Highway Administration, 2017 National Household Travel Survey.

With respect to currently available *Level 1 and Level 2* automation technologies and *Level 3* technologies under development, consumers and other users should understand what the technology is and is not capable of, when human monitoring of the system is needed, and where it should be operated (i.e., appropriate ODD). The private sector may need to consider new approaches for providing information so that consumers can use the technology safely and effectively. As part of their education and training programs and before consumer release, automated vehicle dealers and distributors may want to consider including an on-road or on-track experience demonstrating automated vehicle operations and how humans interact with vehicle controls. Other innovative approaches (e.g., virtual reality (VR) or onboard vehicle systems) may also be considered, tested, and employed.

Public education challenges are different for automated vehicle technologies at higher levels of automation or *Level 4 and Level 5* systems, where the consumer becomes a passenger rather than a driver. For these systems, the members of the public may require more general information and awareness of what the technology is and how they should interact with it, either as passengers or as others sharing the road with automated vehicles.

Developers of automated vehicle technologies are encouraged to develop, document, and maintain employee, dealer, distributor, and consumer education and training programs to address the anticipated differences in the use and operation of automated vehicles from those of the conventional vehicles that the public owns and operates today. Successful programs will provide target users with the necessary level of understanding to utilize these technologies properly, efficiently, and in the safest manner possible.

Consider All Possible Surface Transportation Conditions and Different Roadway Landscapes

Entities that are testing and operating on public roadways will want to consider the whole roadway environment, which could include different infrastructure conditions and operating rules. It will be important to account for all possible surface transportation conditions an ADS may encounter within its ODD. Such conditions, when appropriate, include maneuvering at-grade rail crossings, roundabouts, bicycle lanes, pedestrian walkways and special designated traffic lanes or crossing areas, entrances and driveways, and other potential hazards, especially in different roadway landscapes (e.g., urban versus rural). As part of their important role in the safety assurance of ADS-equipped vehicles, entities are also encouraged to consider such conditions in the design, testing, and validation of the designated fallback method. Entities are encouraged to engage with the U.S. DOT and infrastructure owners and operators to understand the full ODD for safe and efficient operations of automated vehicles.

Work with All Potential User Groups to Incorporate Universal Design Principles

The potential for automation to improve mobility for all Americans is immense, but if products and technologies are not designed with usability by a broad spectrum of travelers in mind, it may not be achieved.

U.S. DOT encourages developers and deployers to work proactively with the disability community to support efforts that focus on the array of accommodations needed for different types of disabilities, and ways to improve mobility as a whole—not just from curb to curb, but also from door to door.

Anticipate Human Factors and Driver Engagement Issues

Consider human factors design for surface transportation—at all levels of automation—for all road users. Safety risks, such as driver distraction and confusion, should influence early stages of design and vehicle development. User-interface usability and comprehension need to be explored, particularly during emergency situations, and in maintaining safety if vehicle functions are compromised.

In addition, it will be important to recognize human factors challenges related to driver awareness and engagement. Entities could consider methods that ensure driver awareness and engagement during ADS-equipped vehicle testing, to mitigate the potential for distraction, fatigue, and other possible risks.

Testing on public roadways is necessary for vehicle automation development and deployment. Public trust can be built during testing by using an in-vehicle driver engagement monitoring system, a second test driver, or other methods. It can be helpful for entities developing ADS technologies to share information with Federal agencies and appropriate organizations about the testing of user interface technologies and designs.

Identify Opportunities for Voluntary Data Exchanges

Voluntary data exchanges can help improve the safety and operations of ADS and lead to the development of industry best practices, voluntary standards, and other useful tools.

Work Zone Data Exchanges

The Work Zone Data Exchange project responds to priorities identified by public and private sector stakeholders. The goal is to develop a harmonized specification for work zone data that infrastructure owners and operators can make available as open feeds that automated vehicles and others can use.

Accurate and up-to-date information about dynamic conditions occurring on the roads—such as work zones—can help automated vehicles navigate safely and efficiently. Many infrastructure owners and operators maintain data on work zone activity, but a common specification for this type of data does not currently exist. This makes it difficult and costly for third parties—including vehicle manufacturers and makers of navigation applications—to access and use work zone data across various jurisdictions.

Several State DOT agencies and private companies are voluntarily participating in the project, with U.S. DOT acting as a technical facilitator. U.S. DOT has been working with these partners to help define the core data elements that should be included in an initial work zone specification and to determine what types of technical assistance the data producers will need to implement it, expand it over time, and address broader work zone data management challenges.

In U.S. DOT's Guiding Principles on Data for Automated Vehicle Safety, available at www.transportation.gov/av/data, the Department defines an approach that seeks to prioritize and enable voluntary data exchanges to address critical issues that could slow the safe integration of ADS technologies. These principles include:

- Promote proactive, data-driven safety, cybersecurity, and privacy-protection practices.
- Act as a facilitator to inspire and enable voluntary data exchanges.
- Start small to demonstrate value, and scale what works toward a larger vision.
- Coordinate across modes to reduce costs, reduce industry burden, and accelerate action.

The industry as a whole should consider working with Federal, State, and local agencies as well as relevant standards bodies (IEEE, SAE International, etc.) to identify opportunities to establish voluntary exchanges of data that can provide mutual benefit and help accelerate the safe integration of automation into the surface transportation system. This can include exchanges of data between the public and private sector regarding infrastructure conditions as well as exchanges among private sector entities to enable mutual learning and risk mitigation.

Any exchanges of data should respect consumer privacy[43] as well as proprietary and confidential business information.

Contribute to the Development of Voluntary, Consensus-Based, and Performance-Oriented Technical Standards

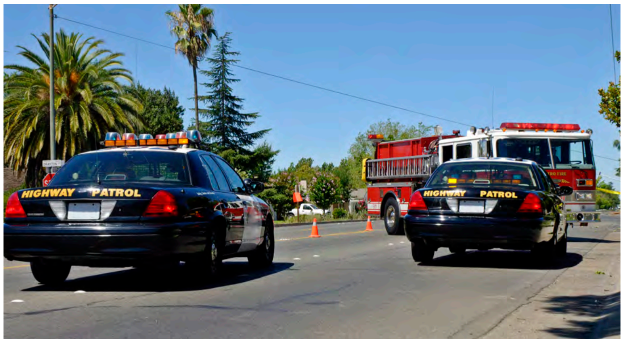
Voluntary standards offer flexibility and responsiveness to the rapid pace of innovation, can encourage investment and bring cost-effective innovation to the market more quickly, and may be validated by private sector conformity assessment and testing protocols. There are existing processes followed by Standards Development Organizations (SDOs), such as SAE International or IEEE, where industry participates in the development of voluntary standards. Industry and SDOs can continue to provide leadership in this area and collaborate with each other, as well as with U.S. DOT and other stakeholders, to address key issues. Areas where industry can support standards development include—but are not limited to—topics such as definitions, taxonomy, testing, interoperability, and performance characteristic definitions.

The Department supports the development and continuing evolution of stakeholder-driven voluntary standards, which in many cases can be an effective non-regulatory means to support interoperable integration of technologies into the transportation system. The Department supports these efforts through multiple mechanisms, including cooperation and funding support to SDOs; cooperation with industry and governmental partners; making Federal, State, and local technical expertise available; and through international coordination.

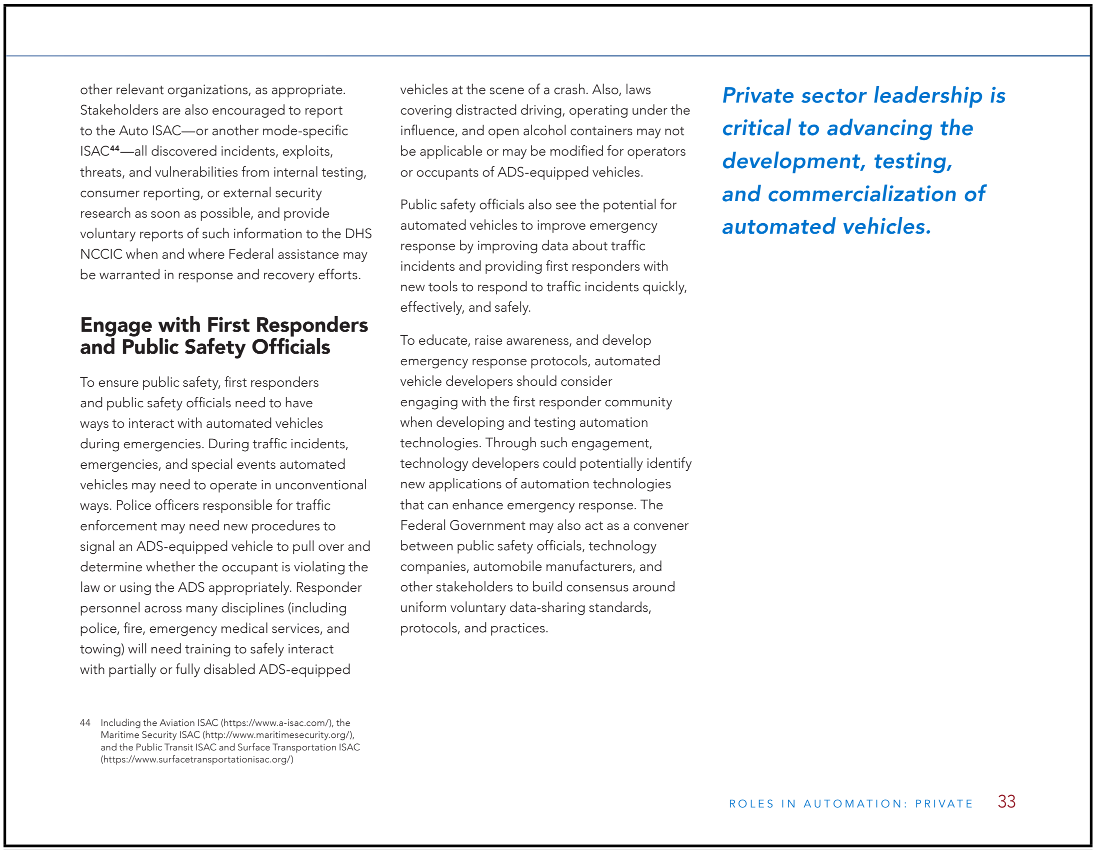
Appendix C provides more information on key topic areas and work underway in standards development for automation.

Adopt Cybersecurity Best Practices

It is the responsibility of ADS developers, vehicle manufacturers, parts suppliers, and all stakeholders who support transportation to follow best practices, and industry standards, for managing cyber risks in the design, integration, testing, and deployment of ADS. As documented in *A Vision for Safety 2.0,* these entities are encouraged to consider and incorporate voluntary guidance, best practices, and design principles published by NIST, NHTSA, SAE International, the Alliance of Automobile Manufacturers, the Association of Global Automakers, the Auto ISAC, and



[43] The Federal Trade Commission maintains oversight over, and provides resources related to, protecting consumer privacy. Additional information is available at https://www.ftc.gov/news-events/media-resources/protecting-consumer-privacy



other relevant organizations, as appropriate. Stakeholders are also encouraged to report to the Auto ISAC—or another mode-specific ISAC[44]—all discovered incidents, exploits, threats, and vulnerabilities from internal testing, consumer reporting, or external security research as soon as possible, and provide voluntary reports of such information to the DHS NCCIC when and where Federal assistance may be warranted in response and recovery efforts.

Engage with First Responders and Public Safety Officials

To ensure public safety, first responders and public safety officials need to have ways to interact with automated vehicles during emergencies. During traffic incidents, emergencies, and special events automated vehicles may need to operate in unconventional ways. Police officers responsible for traffic enforcement may need new procedures to signal an ADS-equipped vehicle to pull over and determine whether the occupant is violating the law or using the ADS appropriately. Responder personnel across many disciplines (including police, fire, emergency medical services, and towing) will need training to safely interact with partially or fully disabled ADS-equipped vehicles at the scene of a crash. Also, laws covering distracted driving, operating under the influence, and open alcohol containers may not be applicable or may be modified for operators or occupants of ADS-equipped vehicles.

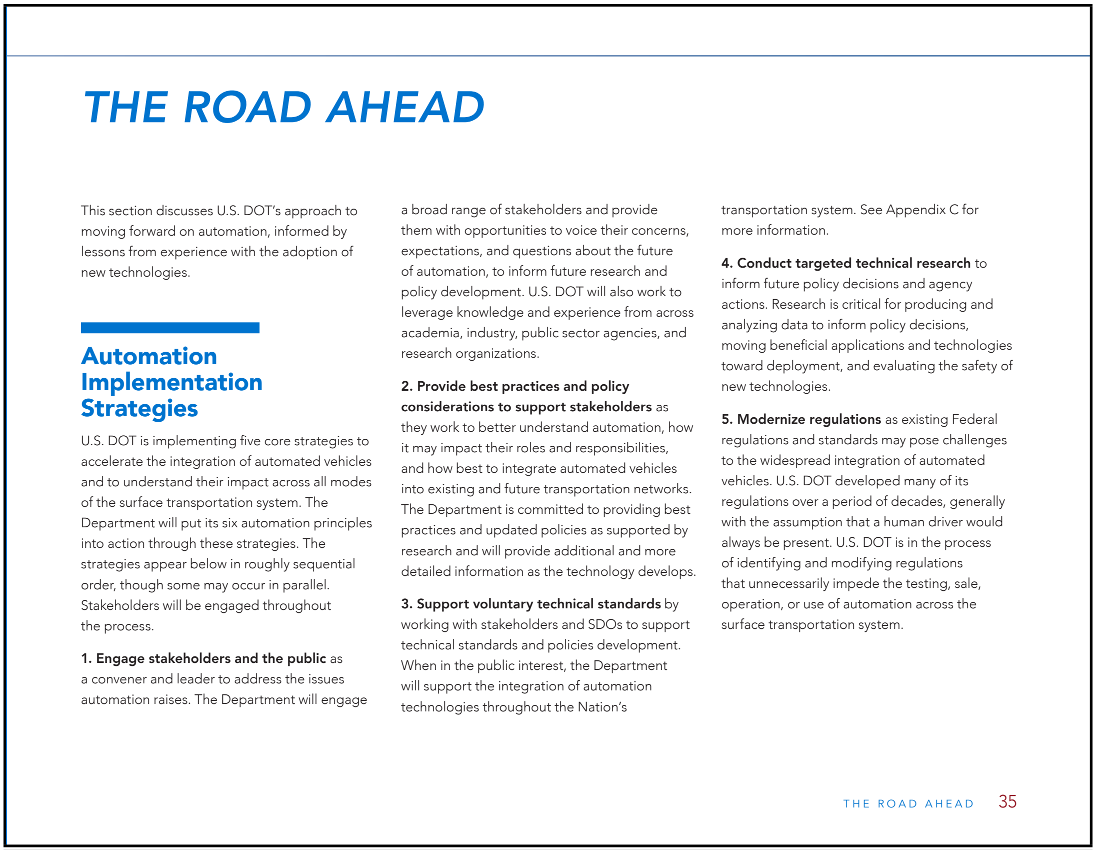
Public safety officials also see the potential for automated vehicles to improve emergency response by improving data about traffic incidents and providing first responders with new tools to respond to traffic incidents quickly, effectively, and safely.

To educate, raise awareness, and develop emergency response protocols, automated vehicle developers should consider engaging with the first responder community when developing and testing automation technologies. Through such engagement, technology developers could potentially identify new applications of automation technologies that can enhance emergency response. The Federal Government may also act as a convener between public safety officials, technology companies, automobile manufacturers, and other stakeholders to build consensus around uniform voluntary data-sharing standards, protocols, and practices.

Private sector leadership is critical to advancing the development, testing, and commercialization of automated vehicles.

[44] Including the Aviation ISAC (https://www.a-isac.com/), the Maritime Security ISAC (http://www.maritimesecurity.org/), and the Public Transit ISAC and Surface Transportation ISAC (https://www.surfacetransportationisac.org/)

U.S. DOT sees a bright future for automation technology and great potential for transforming our surface transportation system for the better, toward a future with enhanced safety, mobility, and economic competitiveness across all transportation modes.



THE ROAD AHEAD

This section discusses U.S. DOT's approach to moving forward on automation, informed by lessons from experience with the adoption of new technologies.

Automation Implementation Strategies

U.S. DOT is implementing five core strategies to accelerate the integration of automated vehicles and to understand their impact across all modes of the surface transportation system. The Department will put its six automation principles into action through these strategies. The strategies appear below in roughly sequential order, though some may occur in parallel. Stakeholders will be engaged throughout the process.

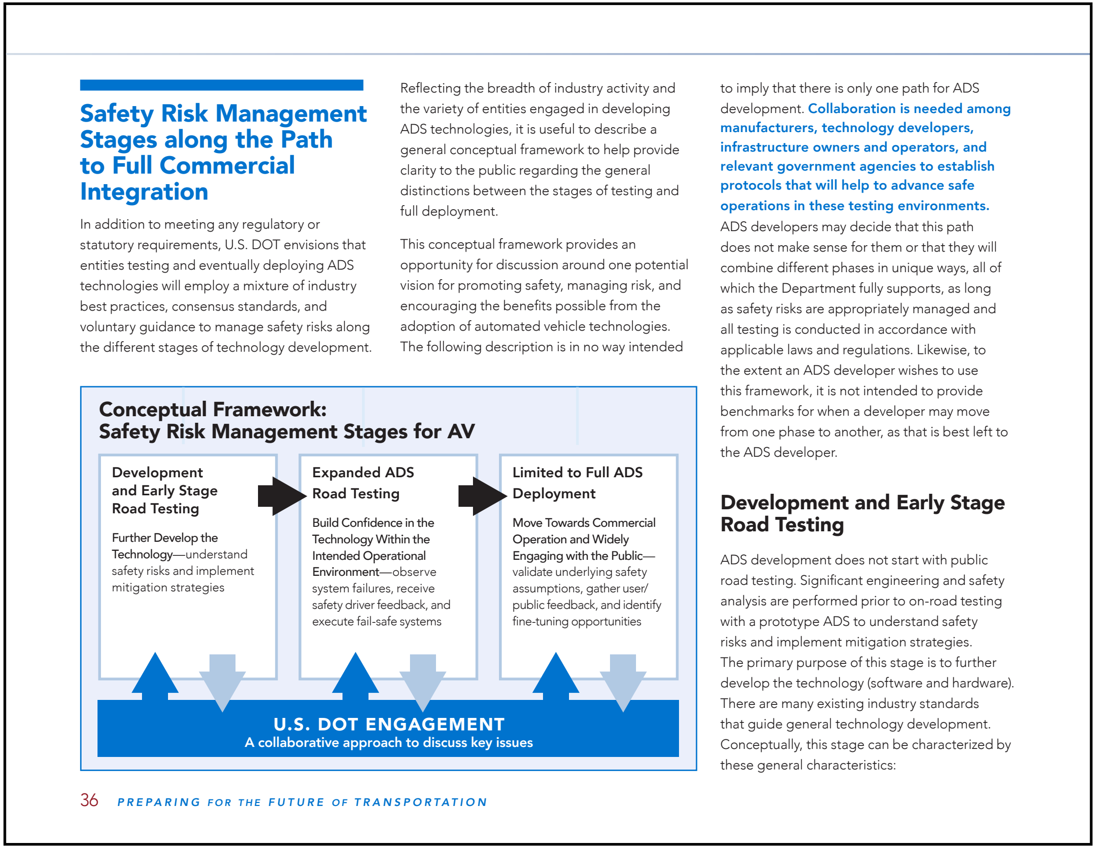
1. Engage stakeholders and the public as a convener and leader to address the issues automation raises. The Department will engage a broad range of stakeholders and provide them with opportunities to voice their concerns, expectations, and questions about the future of automation, to inform future research and policy development. U.S. DOT will also work to leverage knowledge and experience from across academia, industry, public sector agencies, and research organizations.

2. Provide best practices and policy considerations to support stakeholders as they work to better understand automation, how it may impact their roles and responsibilities, and how best to integrate automated vehicles into existing and future transportation networks. The Department is committed to providing best practices and updated policies as supported by research and will provide additional and more detailed information as the technology develops.

3. Support voluntary technical standards by working with stakeholders and SDOs to support technical standards and policies development. When in the public interest, the Department will support the integration of automation technologies throughout the Nation's transportation system. See Appendix C for more information.

4. Conduct targeted technical research to inform future policy decisions and agency actions. Research is critical for producing and analyzing data to inform policy decisions, moving beneficial applications and technologies toward deployment, and evaluating the safety of new technologies.

5. Modernize regulations as existing Federal regulations and standards may pose challenges to the widespread integration of automated vehicles. U.S. DOT developed many of its regulations over a period of decades, generally with the assumption that a human driver would always be present. U.S. DOT is in the process of identifying and modifying regulations that unnecessarily impede the testing, sale, operation, or use of automation across the surface transportation system.



Safety Risk Management Stages along the Path to Full Commercial Integration

In addition to meeting any regulatory or statutory requirements, U.S. DOT envisions that entities testing and eventually deploying ADS technologies will employ a mixture of industry best practices, consensus standards, and voluntary guidance to manage safety risks along the different stages of technology development.

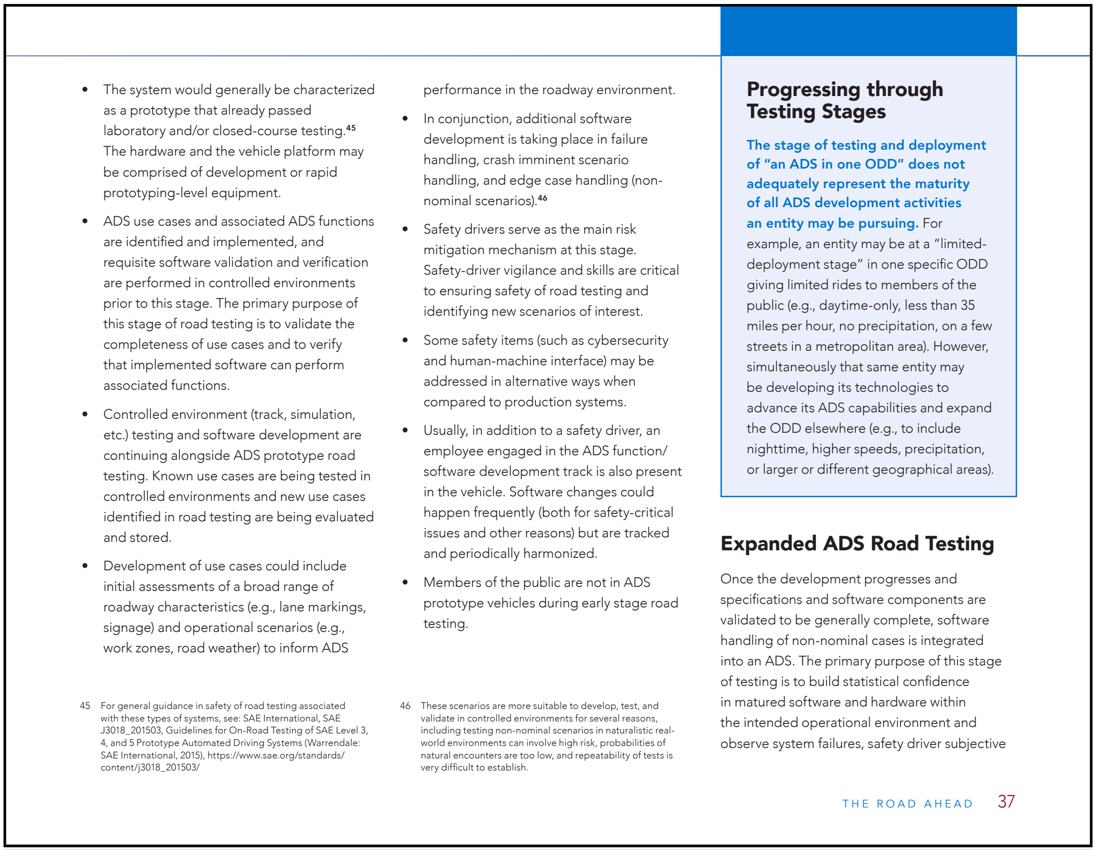
Reflecting the breadth of industry activity and the variety of entities engaged in developing ADS technologies, it is useful to describe a general conceptual framework to help provide clarity to the public regarding the general distinctions between the stages of testing and full deployment.

This conceptual framework provides an opportunity for discussion around one potential vision for promoting safety, managing risk, and encouraging the benefits possible from the adoption of automated vehicle technologies. The following description is in no way intended to imply that there is only one path for ADS development. **Collaboration is needed among manufacturers, technology developers, infrastructure owners and operators, and relevant government agencies to establish protocols that will help to advance safe operations in these testing environments.** ADS developers may decide that this path does not make sense for them or that they will combine different phases in unique ways, all of which the Department fully supports, as long as safety risks are appropriately managed and all testing is conducted in accordance with applicable laws and regulations. Likewise, to the extent an ADS developer wishes to use this framework, it is not intended to provide benchmarks for when a developer may move from one phase to another, as that is best left to the ADS developer.

Conceptual Framework:
Safety Risk Management Stages for AV
Development and Early Stage Road Testing
Further Develop the Technology—understand safety risks and implement mitigation strategies
Expanded ADS Road Testing
Build Confidence in the Technology Within the Intended Operational Environment—observe system failures, receive safety driver feedback, and execute fail-safe systems
Limited to Full ADS Deployment
Move Towards Commercial Operation and Widely Engaging with the Public—validate underlying safety assumptions, gather user/public feedback, and identify fine-tuning opportunities
U.S. DOT ENGAGEMENT
A collaborative approach to discuss key issues

Development and Early Stage Road Testing

ADS development does not start with public road testing. Significant engineering and safety analysis are performed prior to on-road testing with a prototype ADS to understand safety risks and implement mitigation strategies. The primary purpose of this stage is to further develop the technology (software and hardware). There are many existing industry standards that guide general technology development. Conceptually, this stage can be characterized by these general characteristics:



- The system would generally be characterized as a prototype that already passed laboratory and/or closed-course testing.[45] The hardware and the vehicle platform may be comprised of development or rapid prototyping-level equipment.
- ADS use cases and associated ADS functions are identified and implemented, and requisite software validation and verification are performed in controlled environments prior to this stage. The primary purpose of this stage of road testing is to validate the completeness of use cases and to verify that implemented software can perform associated functions.
- Controlled environment (track, simulation, etc.) testing and software development are continuing alongside ADS prototype road testing. Known use cases are being tested in controlled environments and new use cases identified in road testing are being evaluated and stored.
- Development of use cases could include initial assessments of a broad range of roadway characteristics (e.g., lane markings, signage) and operational scenarios (e.g., work zones, road weather) to inform ADS performance in the roadway environment.
- In conjunction, additional software development is taking place in failure handling, crash imminent scenario handling, and edge case handling (non-nominal scenarios).[46]
- Safety drivers serve as the main risk mitigation mechanism at this stage. Safety-driver vigilance and skills are critical to ensuring safety of road testing and identifying new scenarios of interest.
- Some safety items (such as cybersecurity and human-machine interface) may be addressed in alternative ways when compared to production systems.
- Usually, in addition to a safety driver, an employee engaged in the ADS function/software development track is also present in the vehicle. Software changes could happen frequently (both for safety-critical issues and other reasons) but are tracked and periodically harmonized.
- Members of the public are not in ADS prototype vehicles during early stage road testing.

Progressing through Testing Stages

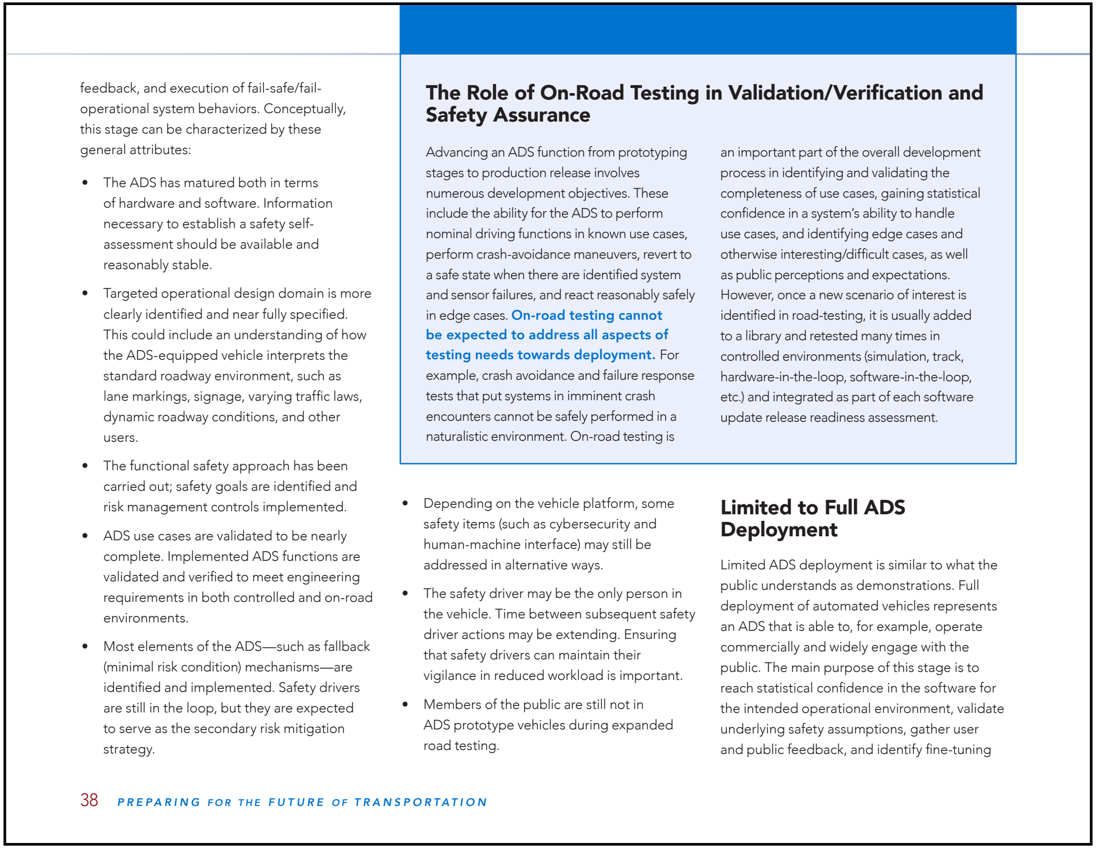
The stage of testing and deployment of "an ADS in one ODD" does not adequately represent the maturity of all ADS development activities an entity may be pursuing. For example, an entity may be at a "limited-deployment stage" in one specific ODD giving limited rides to members of the public (e.g., daytime-only, less than 35 miles per hour, no precipitation, on a few streets in a metropolitan area). However, simultaneously that same entity may be developing its technologies to advance its ADS capabilities and expand the ODD elsewhere (e.g., to include nighttime, higher speeds, precipitation, or larger or different geographical areas).

Expanded ADS Road Testing

Once the development progresses and specifications and software components are validated to be generally complete, software handling of non-nominal cases is integrated into an ADS. The primary purpose of this stage of testing is to build statistical confidence in matured software and hardware within the intended operational environment and observe system failures, safety driver subjective

[45] For general guidance in safety of road testing associated with these types of systems, see: SAE International, SAE J3018_201503, Guidelines for On-Road Testing of SAE Level 3, 4, and 5 Prototype Automated Driving Systems (Warrendale: SAE International, 2015), https://www.sae.org/standards/content/j3018_201503/

[46] These scenarios are more suitable to develop, test, and validate in controlled environments for several reasons, including testing non-nominal scenarios in naturalistic real-world environments can involve high risk, probabilities of natural encounters are too low, and repeatability of tests is very difficult to establish.



feedback, and execution of fail-safe/fail-operational system behaviors. Conceptually, this stage can be characterized by these general attributes:

- The ADS has matured both in terms of hardware and software. Information necessary to establish a safety self-assessment should be available and reasonably stable.
- Targeted operational design domain is more clearly identified and near fully specified. This could include an understanding of how the ADS-equipped vehicle interprets the standard roadway environment, such as lane markings, signage, varying traffic laws, dynamic roadway conditions, and other users.
- The functional safety approach has been carried out; safety goals are identified and risk management controls implemented.
- ADS use cases are validated to be nearly complete. Implemented ADS functions are validated and verified to meet engineering requirements in both controlled and on-road environments.
- Most elements of the ADS—such as fallback (minimal risk condition) mechanisms—are identified and implemented. Safety drivers are still in the loop, but they are expected to serve as the secondary risk mitigation strategy.

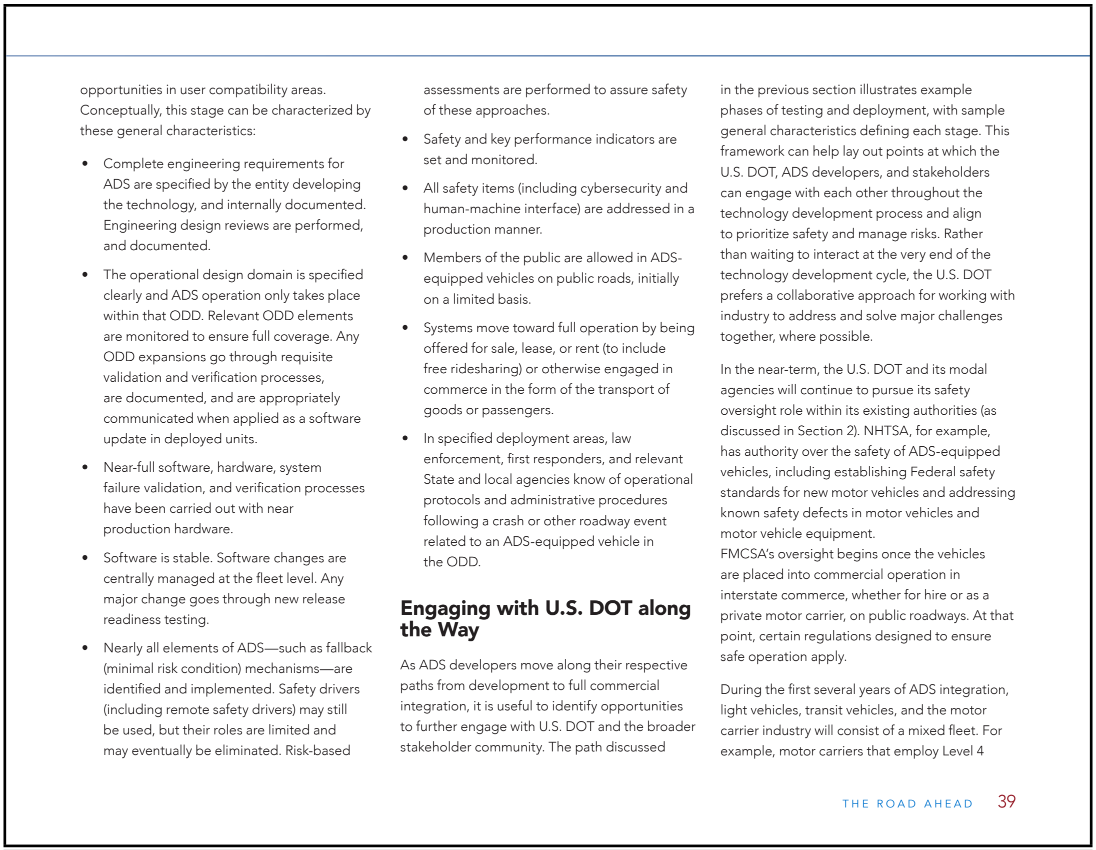
The Role of On-Road Testing in Validation/Verification and Safety Assurance

Advancing an ADS function from prototyping stages to production release involves numerous development objectives. These include the ability for the ADS to perform nominal driving functions in known use cases, perform crash-avoidance maneuvers, revert to a safe state when there are identified system and sensor failures, and react reasonably safely in edge cases. **On-road testing cannot be expected to address all aspects of testing needs towards deployment.** For example, crash avoidance and failure response tests that put systems in imminent crash encounters cannot be safely performed in a naturalistic environment. On-road testing is an important part of the overall development process in identifying and validating the completeness of use cases, gaining statistical confidence in a system's ability to handle use cases, and identifying edge cases and otherwise interesting/difficult cases, as well as public perceptions and expectations. However, once a new scenario of interest is identified in road-testing, it is usually added to a library and retested many times in controlled environments (simulation, track, hardware-in-the-loop, software-in-the-loop, etc.) and integrated as part of each software update release readiness assessment.

- Depending on the vehicle platform, some safety items (such as cybersecurity and human-machine interface) may still be addressed in alternative ways.
- The safety driver may be the only person in the vehicle. Time between subsequent safety driver actions may be extending. Ensuring that safety drivers can maintain their vigilance in reduced workload is important.
- Members of the public are still not in ADS prototype vehicles during expanded road testing.

Limited to Full ADS Deployment

Limited ADS deployment is similar to what the public understands as demonstrations. Full deployment of automated vehicles represents an ADS that is able to, for example, operate commercially and widely engage with the public. The main purpose of this stage is to reach statistical confidence in the software for the intended operational environment, validate underlying safety assumptions, gather user and public feedback, and identify fine-tuning



opportunities in user compatibility areas. Conceptually, this stage can be characterized by these general characteristics:

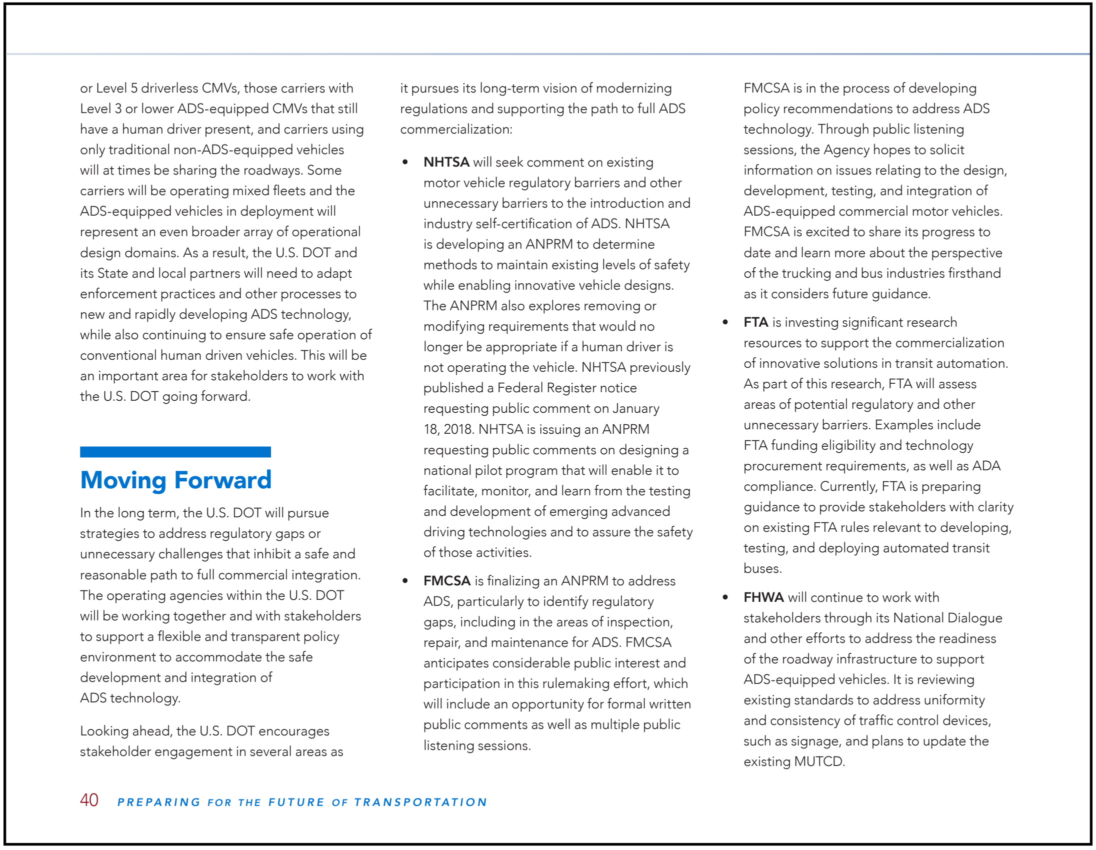
- Complete engineering requirements for ADS are specified by the entity developing the technology, and internally documented. Engineering design reviews are performed, and documented.
- The operational design domain is specified clearly and ADS operation only takes place within that ODD. Relevant ODD elements are monitored to ensure full coverage. Any ODD expansions go through requisite validation and verification processes, are documented, and are appropriately communicated when applied as a software update in deployed units.
- Near-full software, hardware, system failure validation, and verification processes have been carried out with near production hardware.
- Software is stable. Software changes are centrally managed at the fleet level. Any major change goes through new release readiness testing.
- Nearly all elements of ADS—such as fallback (minimal risk condition) mechanisms—are identified and implemented. Safety drivers (including remote safety drivers) may still be used, but their roles are limited and may eventually be eliminated. Risk-based assessments are performed to assure safety of these approaches.
- Safety and key performance indicators are set and monitored.
- All safety items (including cybersecurity and human-machine interface) are addressed in a production manner.
- Members of the public are allowed in ADS-equipped vehicles on public roads, initially on a limited basis.
- Systems move toward full operation by being offered for sale, lease, or rent (to include free ridesharing) or otherwise engaged in commerce in the form of the transport of goods or passengers.
- In specified deployment areas, law enforcement, first responders, and relevant State and local agencies know of operational protocols and administrative procedures following a crash or other roadway event related to an ADS-equipped vehicle in the ODD.

Engaging with U.S. DOT along the Way

As ADS developers move along their respective paths from development to full commercial integration, it is useful to identify opportunities to further engage with U.S. DOT and the broader stakeholder community. The path discussed in the previous section illustrates example phases of testing and deployment, with sample general characteristics defining each stage. This framework can help lay out points at which the U.S. DOT, ADS developers, and stakeholders can engage with each other throughout the technology development process and align to prioritize safety and manage risks. Rather than waiting to interact at the very end of the technology development cycle, the U.S. DOT prefers a collaborative approach for working with industry to address and solve major challenges together, where possible.

In the near-term, the U.S. DOT and its modal agencies will continue to pursue its safety oversight role within its existing authorities (as discussed in Section 2). NHTSA, for example, has authority over the safety of ADS-equipped vehicles, including establishing Federal safety standards for new motor vehicles and addressing known safety defects in motor vehicles and motor vehicle equipment.
FMCSA's oversight begins once the vehicles are placed into commercial operation in interstate commerce, whether for hire or as a private motor carrier, on public roadways. At that point, certain regulations designed to ensure safe operation apply.

During the first several years of ADS integration, light vehicles, transit vehicles, and the motor carrier industry will consist of a mixed fleet. For example, motor carriers that employ Level 4



or Level 5 driverless CMVs, those carriers with Level 3 or lower ADS-equipped CMVs that still have a human driver present, and carriers using only traditional non-ADS-equipped vehicles will at times be sharing the roadways. Some carriers will be operating mixed fleets and the ADS-equipped vehicles in deployment will represent an even broader array of operational design domains. As a result, the U.S. DOT and its State and local partners will need to adapt enforcement practices and other processes to new and rapidly developing ADS technology, while also continuing to ensure safe operation of conventional human driven vehicles. This will be an important area for stakeholders to work with the U.S. DOT going forward.

Moving Forward

In the long term, the U.S. DOT will pursue strategies to address regulatory gaps or unnecessary challenges that inhibit a safe and reasonable path to full commercial integration. The operating agencies within the U.S. DOT will be working together and with stakeholders to support a flexible and transparent policy environment to accommodate the safe development and integration of ADS technology.

Looking ahead, the U.S. DOT encourages stakeholder engagement in several areas as it pursues its long-term vision of modernizing regulations and supporting the path to full ADS commercialization:

- **NHTSA** will seek comment on existing motor vehicle regulatory barriers and other unnecessary barriers to the introduction and industry self-certification of ADS. NHTSA is developing an ANPRM to determine methods to maintain existing levels of safety while enabling innovative vehicle designs. The ANPRM also explores removing or modifying requirements that would no longer be appropriate if a human driver is not operating the vehicle. NHTSA previously published a Federal Register notice requesting public comment on January 18, 2018. NHTSA is issuing an ANPRM requesting public comments on designing a national pilot program that will enable it to facilitate, monitor, and learn from the testing and development of emerging advanced driving technologies and to assure the safety of those activities.
- **FMCSA** is finalizing an ANPRM to address ADS, particularly to identify regulatory gaps, including in the areas of inspection, repair, and maintenance for ADS. FMCSA anticipates considerable public interest and participation in this rulemaking effort, which will include an opportunity for formal written public comments as well as multiple public listening sessions.

  FMCSA is in the process of developing policy recommendations to address ADS technology. Through public listening sessions, the Agency hopes to solicit information on issues relating to the design, development, testing, and integration of ADS-equipped commercial motor vehicles. FMCSA is excited to share its progress to date and learn more about the perspective of the trucking and bus industries firsthand as it considers future guidance.
- **FTA** is investing significant research resources to support the commercialization of innovative solutions in transit automation. As part of this research, FTA will assess areas of potential regulatory and other unnecessary barriers. Examples include FTA funding eligibility and technology procurement requirements, as well as ADA compliance. Currently, FTA is preparing guidance to provide stakeholders with clarity on existing FTA rules relevant to developing, testing, and deploying automated transit buses.
- **FHWA** will continue to work with stakeholders through its National Dialogue and other efforts to address the readiness of the roadway infrastructure to support ADS-equipped vehicles. It is reviewing existing standards to address uniformity and consistency of traffic control devices, such as signage, and plans to update the existing MUTCD.

Stakeholders are encouraged to engage directly with the Department where and when possible to support collaboration. It will be important to gather information and feedback from the stakeholder community, including ADS developers, commercial motor vehicle carriers, transit agencies, infrastructure owners and operators, the public, and other groups to jointly address key challenges and promote safe technology development and deployment.

Conclusion

Over the past century, motor vehicles have provided tremendous mobility benefits, including widespread access to jobs, goods, and services. They have also helped connect many of the most remote and isolated regions of the country to the larger economy. Along with these benefits, however, have come significant safety risks and other challenges. Motor vehicle crashes remain a leading cause of death in the United States, with an estimated 37,133 lives lost on U.S. roads in 2017. Automation has the potential to improve the safety of our transportation system, improve our quality of life, and enhance mobility for Americans, including those who do not drive today.

Many Americans remain skeptical about the notion that their car could one day be driving itself, rather than being driven by humans. We certainly cannot predict the exact way consumers will choose to interact with these technologies. Therefore, the U.S. DOT will not rush to regulate a nascent and rapidly evolving technology. Instead, **the Department supports an environment where innovation can thrive and the American public can be excited and confident about the future of transportation.** Doing this requires a flexible policy architecture.

With *AV 3.0*, U.S. DOT acknowledges the need to modernize existing regulations and think about new ways to deliver on our mission. The Department will work with partners and stakeholders in government, industry, and the public to provide direction, while also remaining open to learning from their experiences and needs. Wherever possible, U.S. DOT will partner with industry to develop voluntary consensus-based standards and will reserve non-prescriptive, performance-based regulations for when they are necessary. The Department will work to assess and minimize the possible harms and spread the benefits of automation technology across the Nation.

Regarding the integration of automation into professional driving tasks, lessons learned through the aviation industry's experience with the introduction of automated systems may be instructive and inform the development of thoughtful, balanced approaches. These are not perfect comparisons, but are still worth considering (See Learning from the History of

Automation in Aviation). The aviation industry discovered that automation required careful consideration of human factors, but led to improved safety ultimately. This transition also did not result in the elimination of pilot jobs, as some had feared.

Despite the great promise of automation technology, important questions remain. For example, as driving becomes more automated, how can safety be improved? How will people interact with these technologies? What happens when a human vehicle operator switches to or from an automated driving mode? As automated driving technologies develop, how will the Nation's 3.8 million professional drivers be affected? Which regulatory obstacles need to be removed? What opportunities and challenges does automation present for long-range regional planning? Will automation lead to increased urban congestion?

U.S. DOT sees a bright future for automation technology and great potential for transforming our surface transportation system for the better, toward a future with enhanced safety, mobility, and economic competitiveness across all transportation modes.

Learning from the History of Automation in the Aviation Workforce

The aviation industry developed technological solutions to help airline pilots manage factors such as high workload, distractions, and abnormal situations. Innovation at that time eventually led to the introduction of autopilot, autothrottle, flight director, sophisticated alerting systems, and more. In part because of these innovations, the safety record for aviation improved significantly.[47] Early automation technology in aviation performed very simple functions; for example, maintaining a set altitude or heading—comparable to conventional cruise control systems offered on most passenger cars today. Pilots readily accepted these systems because they reduced their workload and were easy to understand.

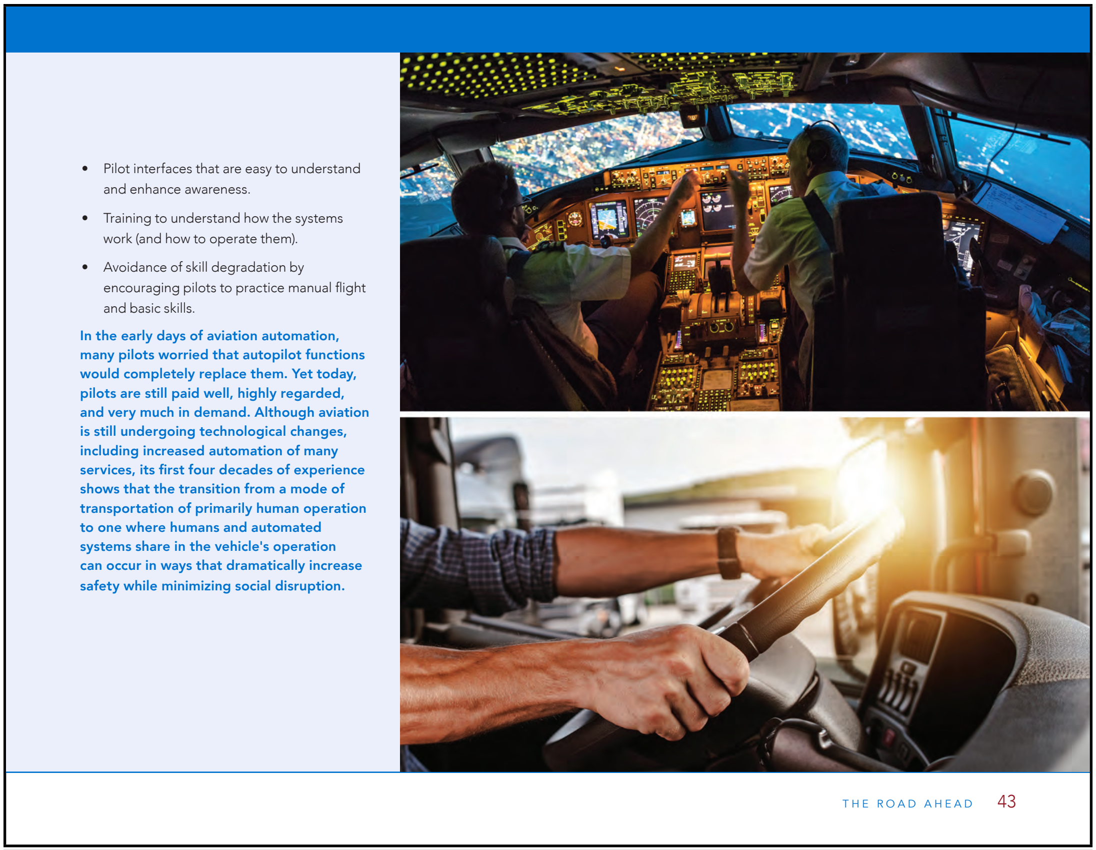
As computer technology became more capable, automation in the flight deck became more complex. For example, it enabled sophisticated navigation using precise flight paths that contributed to more efficient operations. This increased automation came at a cost. It became harder for pilots to understand what the automated systems were doing, yet they remained responsible for taking over when the automated systems reached the limits of their operating domains or malfunctioned. Pilots were also encouraged to use automation to the exclusion of manual flight controls, potentially degrading manual flight skills.

Systems that alert pilots to hazardous conditions (e.g., proximity to the ground or to other aircraft—lane departure alerts are an analogous example offered in many passenger cars) have also contributed significantly to aviation safety despite initial challenges. Early alert systems sometimes had a high number of false alarms, so pilots did not trust them. Many improvements were made, such as better algorithms, better sensors, and improved and standardized display of alerts (and associated information) on the flight deck. These improvements have led to more reliable alerts and pilots are more willing to heed them.

Automation has undeniably made flying safer by supporting pilots. The characteristics that have improved trust in and effectiveness of these systems include:

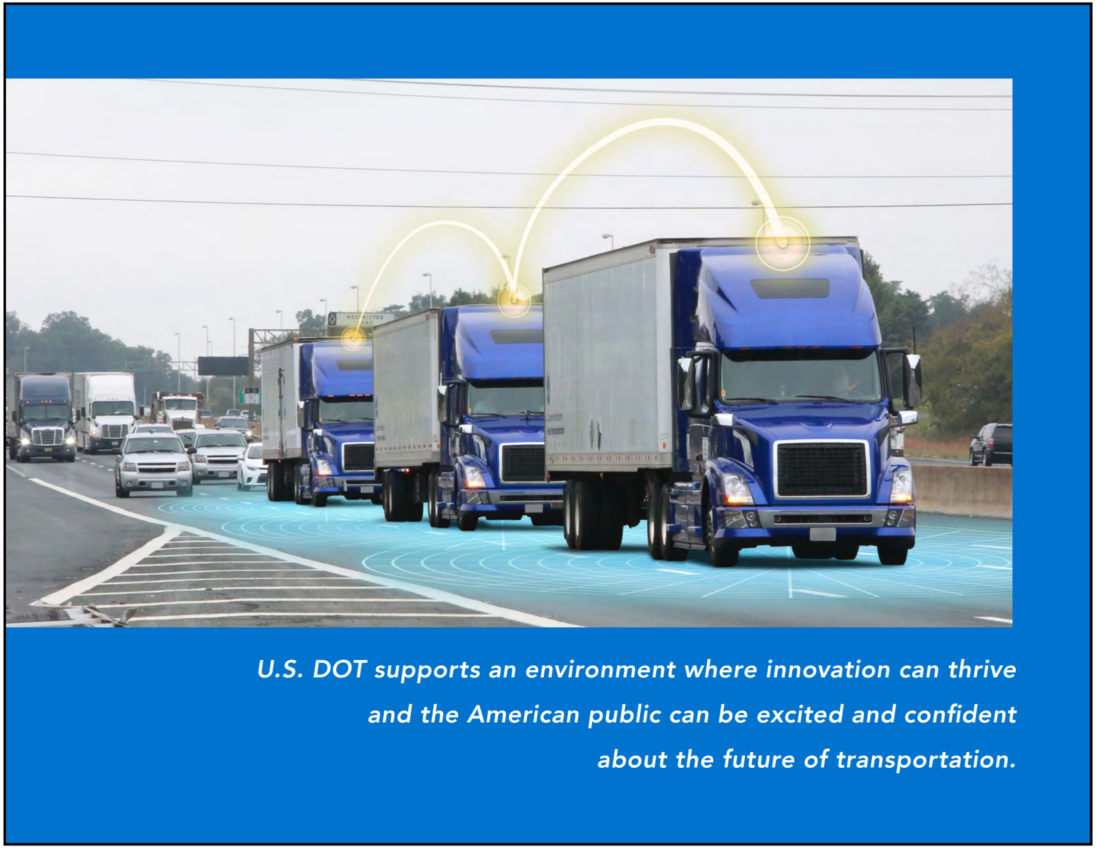
- Reliable, robust systems that minimize false or missed alarms/reports.

[47] Federal Aviation Administration, Operational use of flight path management systems, Final Report, Performance-based operations Aviation Rulemaking Committee/Commercial Aviation Safety Team, Flight Deck Automation Working Group (Washington: Federal Aviation Administration, 2013), https://www.faa.gov/aircraft/air_cert/design_approvals/human_factors/media/OUFPMS_Report.pdf.

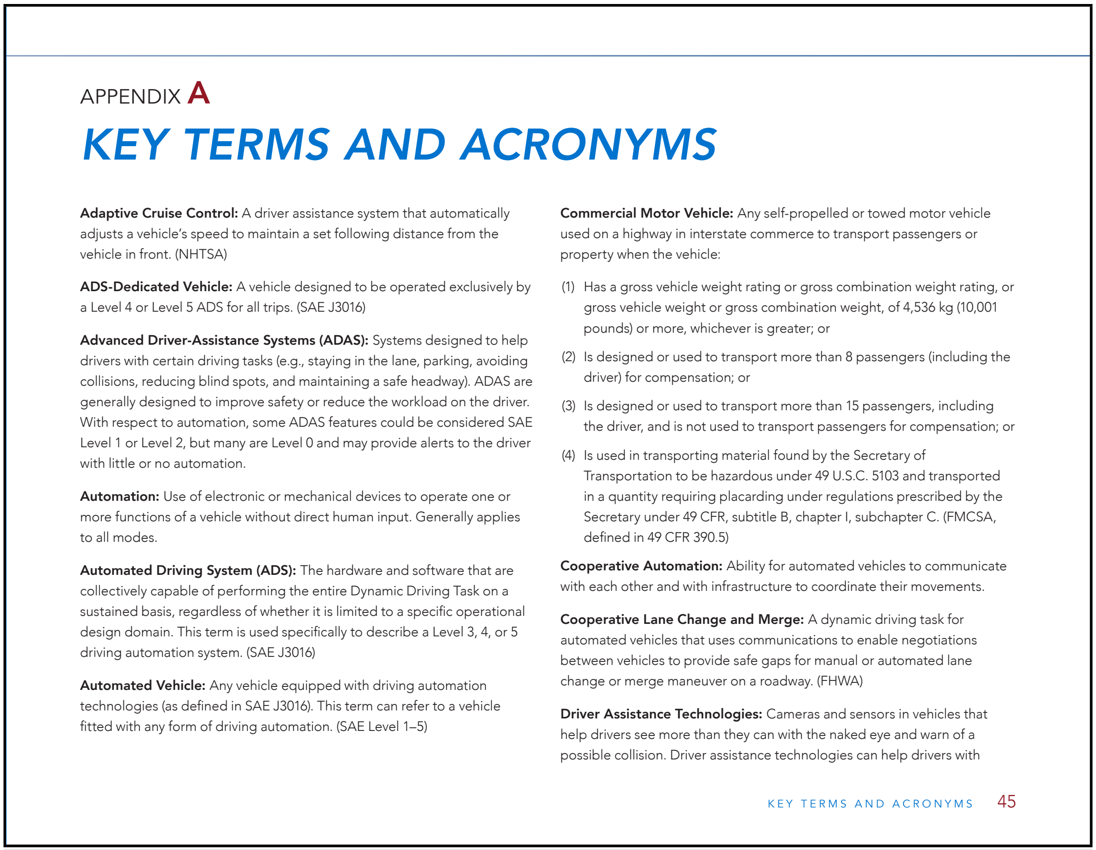


- Pilot interfaces that are easy to understand and enhance awareness.
- Training to understand how the systems work (and how to operate them).
- Avoidance of skill degradation by encouraging pilots to practice manual flight and basic skills.

In the early days of aviation automation, many pilots worried that autopilot functions would completely replace them. Yet today, pilots are still paid well, highly regarded, and very much in demand. Although aviation is still undergoing technological changes, including increased automation of many services, its first four decades of experience shows that the transition from a mode of transportation of primarily human operation to one where humans and automated systems share in the vehicle's operation can occur in ways that dramatically increase safety while minimizing social disruption.



U.S. DOT supports an environment where innovation can thrive and the American public can be excited and confident about the future of transportation.



APPENDIX A

KEY TERMS AND ACRONYMS

Adaptive Cruise Control: A driver assistance system that automatically adjusts a vehicle's speed to maintain a set following distance from the vehicle in front. (NHTSA)

ADS-Dedicated Vehicle: A vehicle designed to be operated exclusively by a Level 4 or Level 5 ADS for all trips. (SAE J3016)

Advanced Driver-Assistance Systems (ADAS): Systems designed to help drivers with certain driving tasks (e.g., staying in the lane, parking, avoiding collisions, reducing blind spots, and maintaining a safe headway). ADAS are generally designed to improve safety or reduce the workload on the driver. With respect to automation, some ADAS features could be considered SAE Level 1 or Level 2, but many are Level 0 and may provide alerts to the driver with little or no automation.

Automation: Use of electronic or mechanical devices to operate one or more functions of a vehicle without direct human input. Generally applies to all modes.

Automated Driving System (ADS): The hardware and software that are collectively capable of performing the entire Dynamic Driving Task on a sustained basis, regardless of whether it is limited to a specific operational design domain. This term is used specifically to describe a Level 3, 4, or 5 driving automation system. (SAE J3016)

Automated Vehicle: Any vehicle equipped with driving automation technologies (as defined in SAE J3016). This term can refer to a vehicle fitted with any form of driving automation. (SAE Level 1–5)

Commercial Motor Vehicle: Any self-propelled or towed motor vehicle used on a highway in interstate commerce to transport passengers or property when the vehicle:

(1) Has a gross vehicle weight rating or gross combination weight rating, or gross vehicle weight or gross combination weight, of 4,536 kg (10,001 pounds) or more, whichever is greater; or

(2) Is designed or used to transport more than 8 passengers (including the driver) for compensation; or

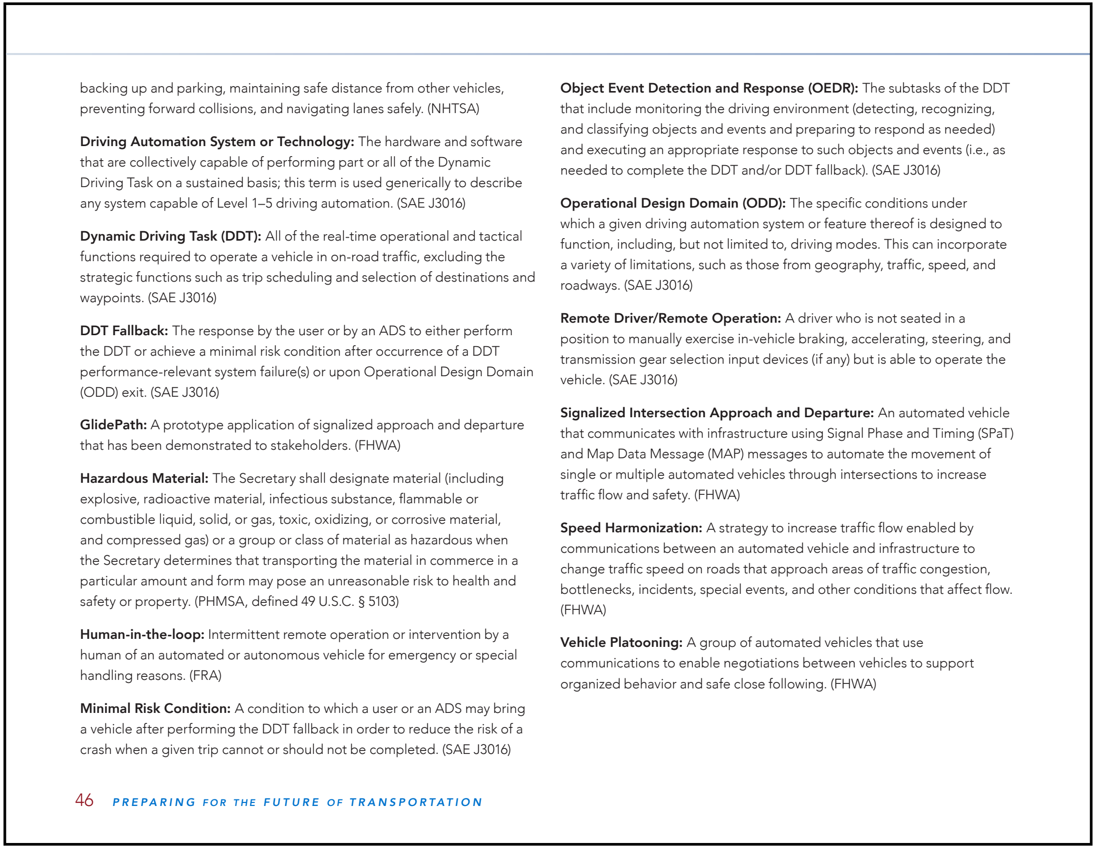
(3) Is designed or used to transport more than 15 passengers, including the driver, and is not used to transport passengers for compensation; or

(4) Is used in transporting material found by the Secretary of Transportation to be hazardous under 49 U.S.C. 5103 and transported in a quantity requiring placarding under regulations prescribed by the Secretary under 49 CFR, subtitle B, chapter I, subchapter C. (FMCSA, defined in 49 CFR 390.5)

Cooperative Automation: Ability for automated vehicles to communicate with each other and with infrastructure to coordinate their movements.

Cooperative Lane Change and Merge: A dynamic driving task for automated vehicles that uses communications to enable negotiations between vehicles to provide safe gaps for manual or automated lane change or merge maneuver on a roadway. (FHWA)

Driver Assistance Technologies: Cameras and sensors in vehicles that help drivers see more than they can with the naked eye and warn of a possible collision. Driver assistance technologies can help drivers with backing up and parking, maintaining safe distance from other vehicles, preventing forward collisions, and navigating lanes safely. (NHTSA)



Driving Automation System or Technology: The hardware and software that are collectively capable of performing part or all of the Dynamic Driving Task on a sustained basis; this term is used generically to describe any system capable of Level 1–5 driving automation. (SAE J3016)

Dynamic Driving Task (DDT): All of the real-time operational and tactical functions required to operate a vehicle in on-road traffic, excluding the strategic functions such as trip scheduling and selection of destinations and waypoints. (SAE J3016)

DDT Fallback: The response by the user or by an ADS to either perform the DDT or achieve a minimal risk condition after occurrence of a DDT performance-relevant system failure(s) or upon Operational Design Domain (ODD) exit. (SAE J3016)

GlidePath: A prototype application of signalized approach and departure that has been demonstrated to stakeholders. (FHWA)

Hazardous Material: The Secretary shall designate material (including explosive, radioactive material, infectious substance, flammable or combustible liquid, solid, or gas, toxic, oxidizing, or corrosive material, and compressed gas) or a group or class of material as hazardous when the Secretary determines that transporting the material in commerce in a particular amount and form may pose an unreasonable risk to health and safety or property. (PHMSA, defined 49 U.S.C. § 5103)

Human-in-the-loop: Intermittent remote operation or intervention by a human of an automated or autonomous vehicle for emergency or special handling reasons. (FRA)

Minimal Risk Condition: A condition to which a user or an ADS may bring a vehicle after performing the DDT fallback in order to reduce the risk of a crash when a given trip cannot or should not be completed. (SAE J3016)

Object Event Detection and Response (OEDR): The subtasks of the DDT that include monitoring the driving environment (detecting, recognizing, and classifying objects and events and preparing to respond as needed) and executing an appropriate response to such objects and events (i.e., as needed to complete the DDT and/or DDT fallback). (SAE J3016)

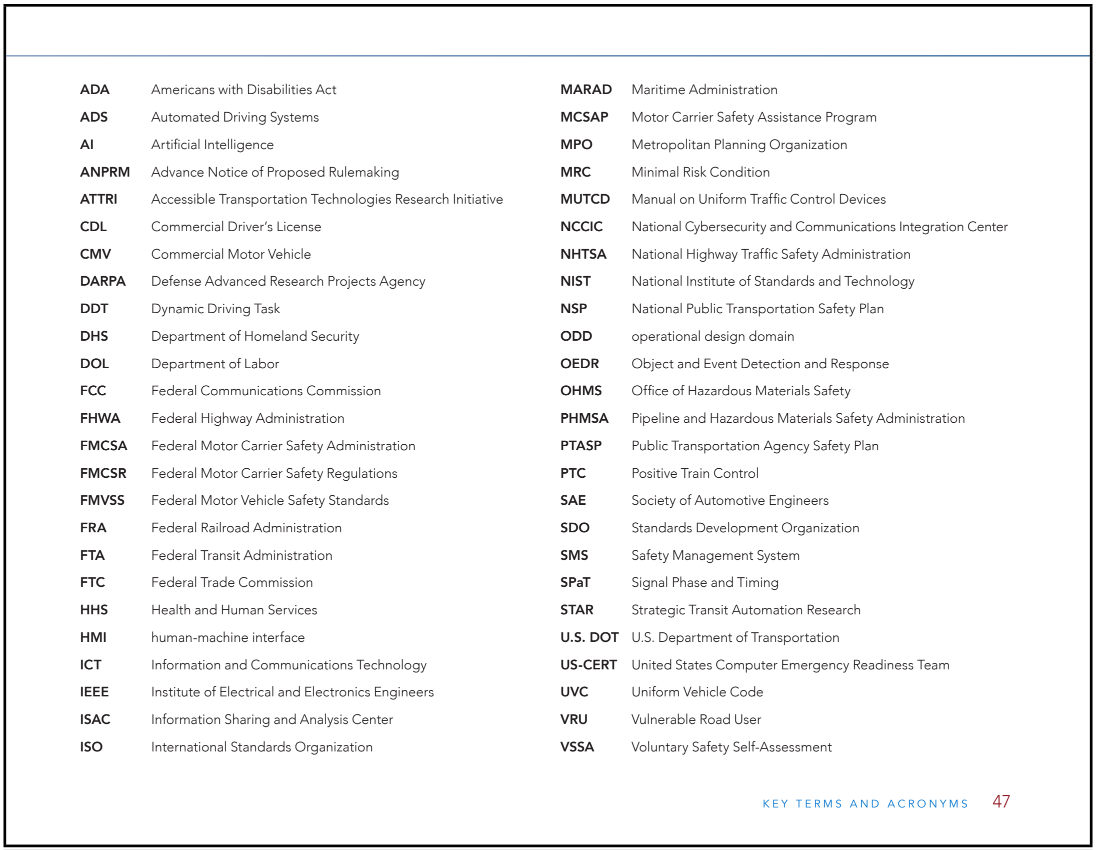
Operational Design Domain (ODD): The specific conditions under which a given driving automation system or feature thereof is designed to function, including, but not limited to, driving modes. This can incorporate a variety of limitations, such as those from geography, traffic, speed, and roadways. (SAE J3016)

Remote Driver/Remote Operation: A driver who is not seated in a position to manually exercise in-vehicle braking, accelerating, steering, and transmission gear selection input devices (if any) but is able to operate the vehicle. (SAE J3016)

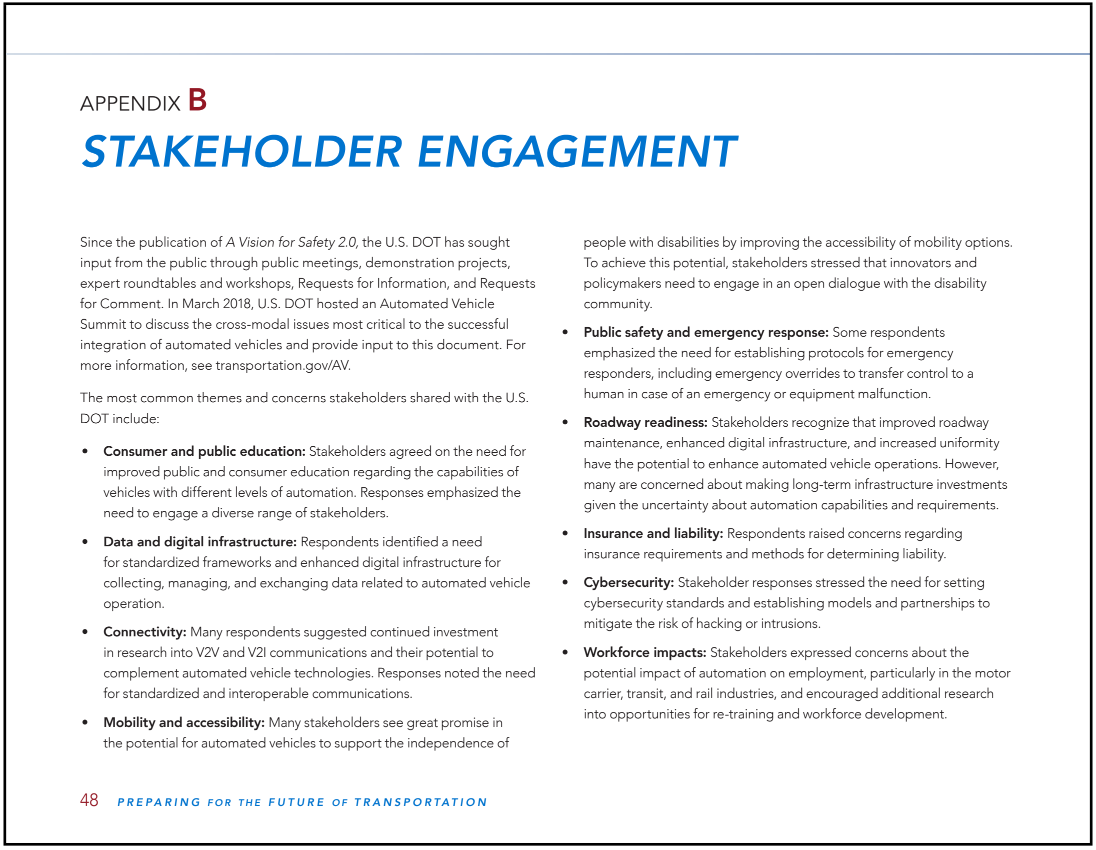
Signalized Intersection Approach and Departure: An automated vehicle that communicates with infrastructure using Signal Phase and Timing (SPaT) and Map Data Message (MAP) messages to automate the movement of single or multiple automated vehicles through intersections to increase traffic flow and safety. (FHWA)

Speed Harmonization: A strategy to increase traffic flow enabled by communications between an automated vehicle and infrastructure to change traffic speed on roads that approach areas of traffic congestion, bottlenecks, incidents, special events, and other conditions that affect flow. (FHWA)

Vehicle Platooning: A group of automated vehicles that use communications to enable negotiations between vehicles to support organized behavior and safe close following. (FHWA)



ADA	Americans with Disabilities Act	**MARAD**	Maritime Administration
ADS	Automated Driving Systems	**MCSAP**	Motor Carrier Safety Assistance Program
AI	Artificial Intelligence	**MPO**	Metropolitan Planning Organization
ANPRM	Advance Notice of Proposed Rulemaking	**MRC**	Minimal Risk Condition
ATTRI	Accessible Transportation Technologies Research Initiative	**MUTCD**	Manual on Uniform Traffic Control Devices
CDL	Commercial Driver's License	**NCCIC**	National Cybersecurity and Communications Integration Center
CMV	Commercial Motor Vehicle	**NHTSA**	National Highway Traffic Safety Administration
DARPA	Defense Advanced Research Projects Agency	**NIST**	National Institute of Standards and Technology
DDT	Dynamic Driving Task	**NSP**	National Public Transportation Safety Plan
DHS	Department of Homeland Security	**ODD**	operational design domain
DOL	Department of Labor	**OEDR**	Object and Event Detection and Response
FCC	Federal Communications Commission	**OHMS**	Office of Hazardous Materials Safety
FHWA	Federal Highway Administration	**PHMSA**	Pipeline and Hazardous Materials Safety Administration
FMCSA	Federal Motor Carrier Safety Administration	**PTASP**	Public Transportation Agency Safety Plan
FMCSR	Federal Motor Carrier Safety Regulations	**PTC**	Positive Train Control
FMVSS	Federal Motor Vehicle Safety Standards	**SAE**	Society of Automotive Engineers
FRA	Federal Railroad Administration	**SDO**	Standards Development Organization
FTA	Federal Transit Administration	**SMS**	Safety Management System
FTC	Federal Trade Commission	**SPaT**	Signal Phase and Timing
HHS	Health and Human Services	**STAR**	Strategic Transit Automation Research
HMI	human-machine interface	**U.S. DOT**	U.S. Department of Transportation
ICT	Information and Communications Technology	**US-CERT**	United States Computer Emergency Readiness Team
IEEE	Institute of Electrical and Electronics Engineers	**UVC**	Uniform Vehicle Code
ISAC	Information Sharing and Analysis Center	**VRU**	Vulnerable Road User
ISO	International Standards Organization	**VSSA**	Voluntary Safety Self-Assessment



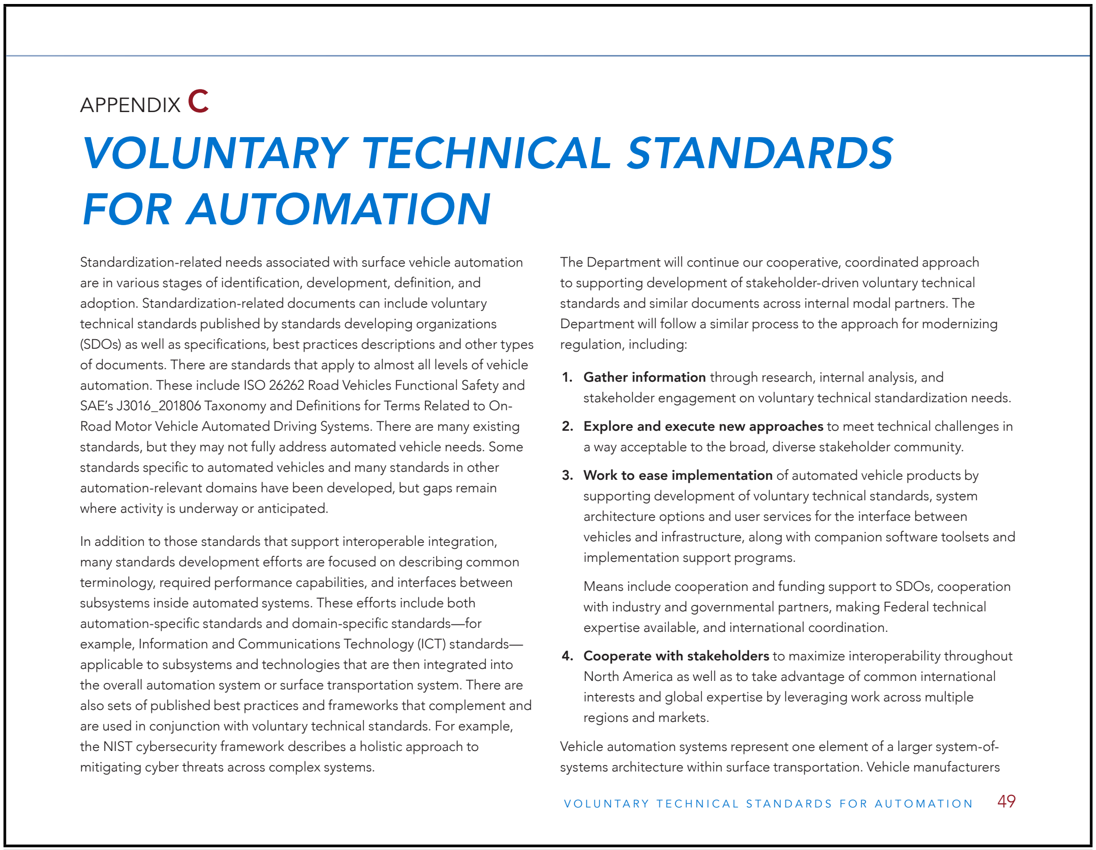
APPENDIX **B**

STAKEHOLDER ENGAGEMENT

Since the publication of *A Vision for Safety 2.0*, the U.S. DOT has sought input from the public through public meetings, demonstration projects, expert roundtables and workshops, Requests for Information, and Requests for Comment. In March 2018, U.S. DOT hosted an Automated Vehicle Summit to discuss the cross-modal issues most critical to the successful integration of automated vehicles and provide input to this document. For more information, see transportation.gov/AV.

The most common themes and concerns stakeholders shared with the U.S. DOT include:

- **Consumer and public education:** Stakeholders agreed on the need for improved public and consumer education regarding the capabilities of vehicles with different levels of automation. Responses emphasized the need to engage a diverse range of stakeholders.
- **Data and digital infrastructure:** Respondents identified a need for standardized frameworks and enhanced digital infrastructure for collecting, managing, and exchanging data related to automated vehicle operation.
- **Connectivity:** Many respondents suggested continued investment in research into V2V and V2I communications and their potential to complement automated vehicle technologies. Responses noted the need for standardized and interoperable communications.
- **Mobility and accessibility:** Many stakeholders see great promise in the potential for automated vehicles to support the independence of people with disabilities by improving the accessibility of mobility options. To achieve this potential, stakeholders stressed that innovators and policymakers need to engage in an open dialogue with the disability community.
- **Public safety and emergency response:** Some respondents emphasized the need for establishing protocols for emergency responders, including emergency overrides to transfer control to a human in case of an emergency or equipment malfunction.
- **Roadway readiness:** Stakeholders recognize that improved roadway maintenance, enhanced digital infrastructure, and increased uniformity have the potential to enhance automated vehicle operations. However, many are concerned about making long-term infrastructure investments given the uncertainty about automation capabilities and requirements.
- **Insurance and liability:** Respondents raised concerns regarding insurance requirements and methods for determining liability.
- **Cybersecurity:** Stakeholder responses stressed the need for setting cybersecurity standards and establishing models and partnerships to mitigate the risk of hacking or intrusions.
- **Workforce impacts:** Stakeholders expressed concerns about the potential impact of automation on employment, particularly in the motor carrier, transit, and rail industries, and encouraged additional research into opportunities for re-training and workforce development.



APPENDIX C

VOLUNTARY TECHNICAL STANDARDS FOR AUTOMATION

Standardization-related needs associated with surface vehicle automation are in various stages of identification, development, definition, and adoption. Standardization-related documents can include voluntary technical standards published by standards developing organizations (SDOs) as well as specifications, best practices descriptions and other types of documents. There are standards that apply to almost all levels of vehicle automation. These include ISO 26262 Road Vehicles Functional Safety and SAE's J3016_201806 Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. There are many existing standards, but they may not fully address automated vehicle needs. Some standards specific to automated vehicles and many standards in other automation-relevant domains have been developed, but gaps remain where activity is underway or anticipated.

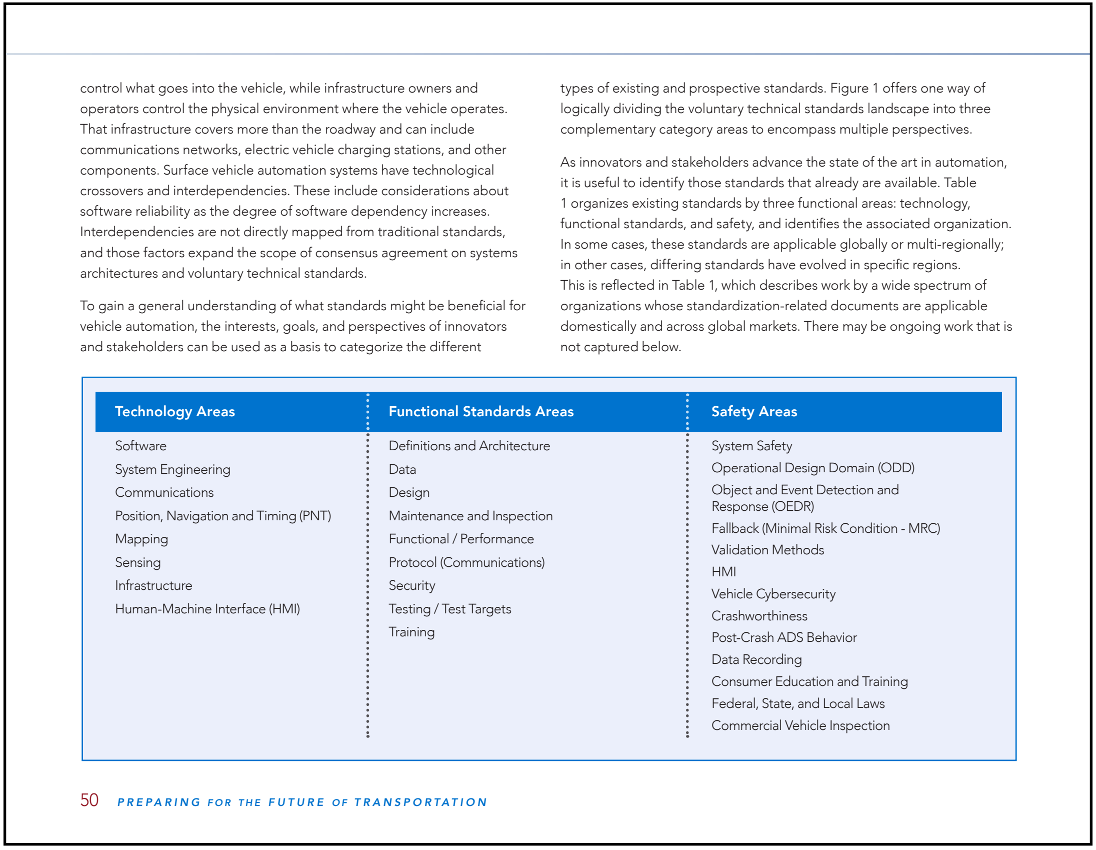
In addition to those standards that support interoperable integration, many standards development efforts are focused on describing common terminology, required performance capabilities, and interfaces between subsystems inside automated systems. These efforts include both automation-specific standards and domain-specific standards—for example, Information and Communications Technology (ICT) standards—applicable to subsystems and technologies that are then integrated into the overall automation system or surface transportation system. There are also sets of published best practices and frameworks that complement and are used in conjunction with voluntary technical standards. For example, the NIST cybersecurity framework describes a holistic approach to mitigating cyber threats across complex systems.

The Department will continue our cooperative, coordinated approach to supporting development of stakeholder-driven voluntary technical standards and similar documents across internal modal partners. The Department will follow a similar process to the approach for modernizing regulation, including:

1. **Gather information** through research, internal analysis, and stakeholder engagement on voluntary technical standardization needs.
2. **Explore and execute new approaches** to meet technical challenges in a way acceptable to the broad, diverse stakeholder community.
3. **Work to ease implementation** of automated vehicle products by supporting development of voluntary technical standards, system architecture options and user services for the interface between vehicles and infrastructure, along with companion software toolsets and implementation support programs.

   Means include cooperation and funding support to SDOs, cooperation with industry and governmental partners, making Federal technical expertise available, and international coordination.
4. **Cooperate with stakeholders** to maximize interoperability throughout North America as well as to take advantage of common international interests and global expertise by leveraging work across multiple regions and markets.

Vehicle automation systems represent one element of a larger system-of-systems architecture within surface transportation. Vehicle manufacturers control what goes into the vehicle, while infrastructure owners and operators control the physical environment where the vehicle operates. That infrastructure covers more than the roadway and can include communications networks, electric vehicle charging stations, and other components. Surface vehicle automation systems have technological crossovers and interdependencies. These include considerations about software reliability as the degree of software dependency increases. Interdependencies are not directly mapped from traditional standards, and those factors expand the scope of consensus agreement on systems architectures and voluntary technical standards.



To gain a general understanding of what standards might be beneficial for vehicle automation, the interests, goals, and perspectives of innovators and stakeholders can be used as a basis to categorize the different types of existing and prospective standards. Figure 1 offers one way of logically dividing the voluntary technical standards landscape into three complementary category areas to encompass multiple perspectives.

As innovators and stakeholders advance the state of the art in automation, it is useful to identify those standards that already are available. Table 1 organizes existing standards by three functional areas: technology, functional standards, and safety, and identifies the associated organization. In some cases, these standards are applicable globally or multi-regionally; in other cases, differing standards have evolved in specific regions. This is reflected in Table 1, which describes work by a wide spectrum of organizations whose standardization-related documents are applicable domestically and across global markets. There may be ongoing work that is not captured below.

Technology Areas	Functional Standards Areas	Safety Areas
Software	Definitions and Architecture	System Safety
System Engineering	Data	Operational Design Domain (ODD)
Communications	Design	Object and Event Detection and Response (OEDR)
Position, Navigation and Timing (PNT)	Maintenance and Inspection	Fallback (Minimal Risk Condition - MRC)
Mapping	Functional / Performance	Validation Methods
Sensing	Protocol (Communications)	HMI
Infrastructure	Security	Vehicle Cybersecurity
Human-Machine Interface (HMI)	Testing / Test Targets	Crashworthiness
	Training	Post-Crash ADS Behavior
		Data Recording
		Consumer Education and Training
		Federal, State, and Local Laws
		Commercial Vehicle Inspection

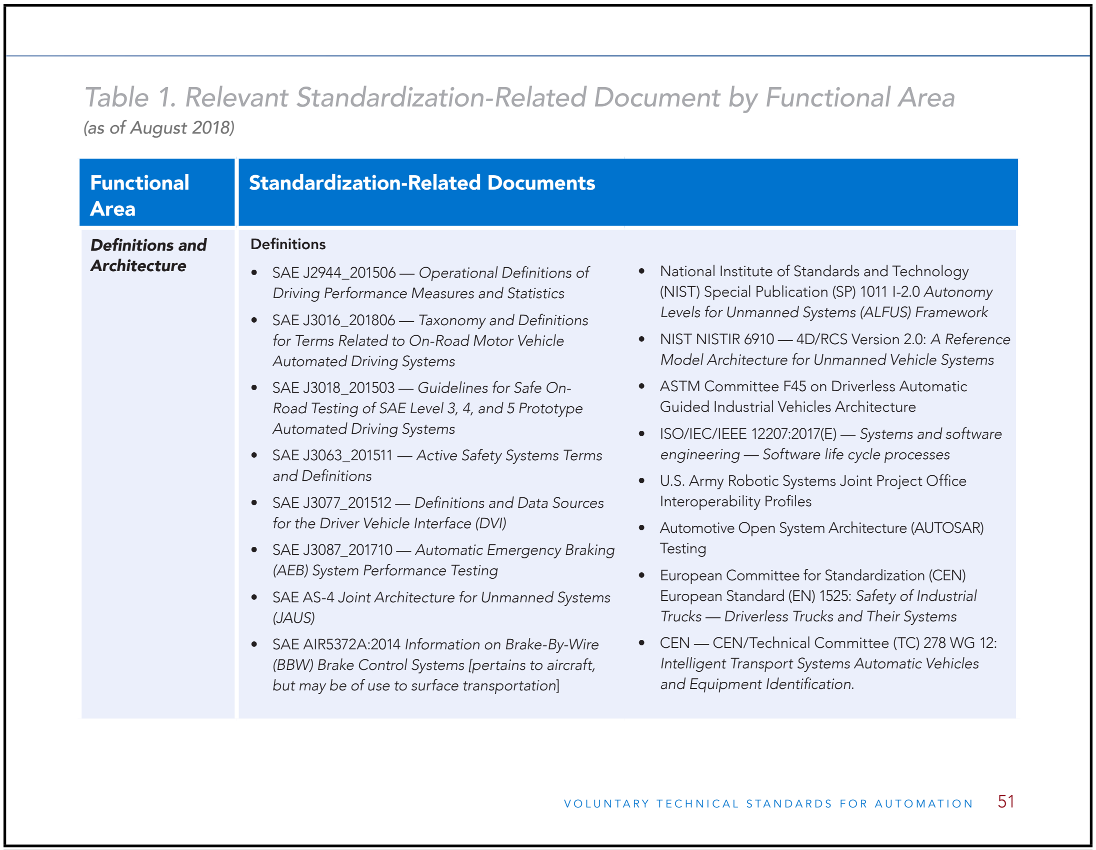
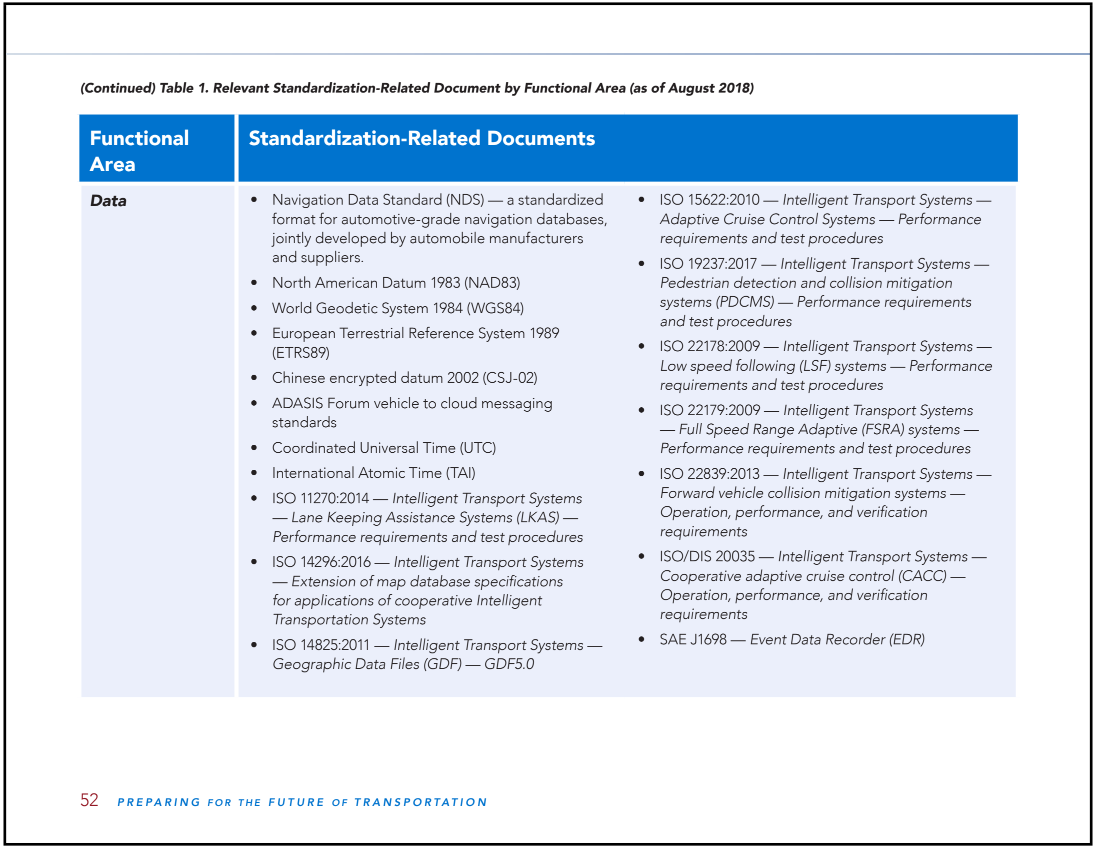


Table 1. Relevant Standardization-Related Document by Functional Area

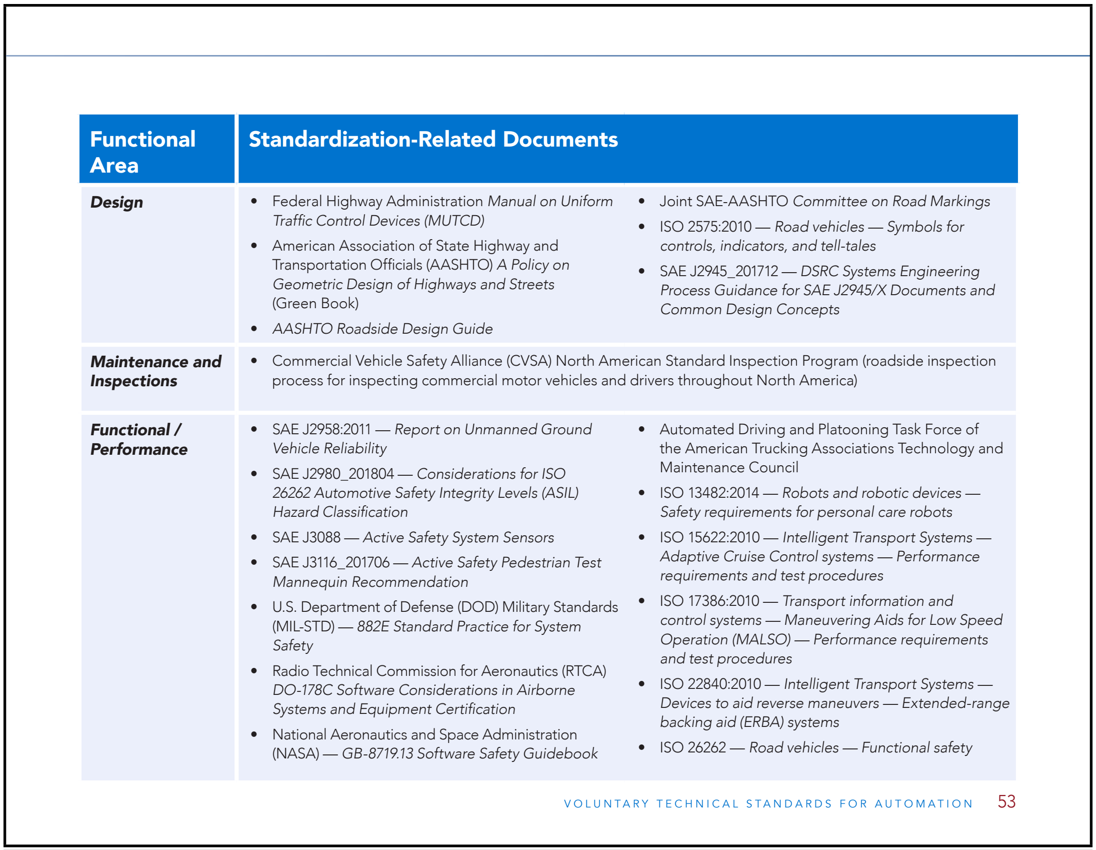
(as of August 2018)

Functional Area	Standardization-Related Documents
Definitions and Architecture	Definitions • SAE J2944_201506 — *Operational Definitions of Driving Performance Measures and Statistics* • SAE J3016_201806 — *Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems* • SAE J3018_201503 — *Guidelines for Safe On-Road Testing of SAE Level 3, 4, and 5 Prototype Automated Driving Systems* • SAE J3063_201511 — *Active Safety Systems Terms and Definitions* • SAE J3077_201512 — *Definitions and Data Sources for the Driver Vehicle Interface (DVI)* • SAE J3087_201710 — *Automatic Emergency Braking (AEB) System Performance Testing* • SAE AS-4 *Joint Architecture for Unmanned Systems (JAUS)* • SAE AIR5372A:2014 *Information on Brake-By-Wire (BBW) Brake Control Systems [pertains to aircraft, but may be of use to surface transportation]* • National Institute of Standards and Technology (NIST) Special Publication (SP) 1011 I-2.0 *Autonomy Levels for Unmanned Systems (ALFUS) Framework* • NIST NISTIR 6910 — 4D/RCS Version 2.0: *A Reference Model Architecture for Unmanned Vehicle Systems* • ASTM Committee F45 on Driverless Automatic Guided Industrial Vehicles Architecture • ISO/IEC/IEEE 12207:2017(E) — *Systems and software engineering — Software life cycle processes* • U.S. Army Robotic Systems Joint Project Office Interoperability Profiles • Automotive Open System Architecture (AUTOSAR) Testing • European Committee for Standardization (CEN) European Standard (EN) 1525: *Safety of Industrial Trucks — Driverless Trucks and Their Systems* • CEN — CEN/Technical Committee (TC) 278 WG 12: *Intelligent Transport Systems Automatic Vehicles and Equipment Identification.*

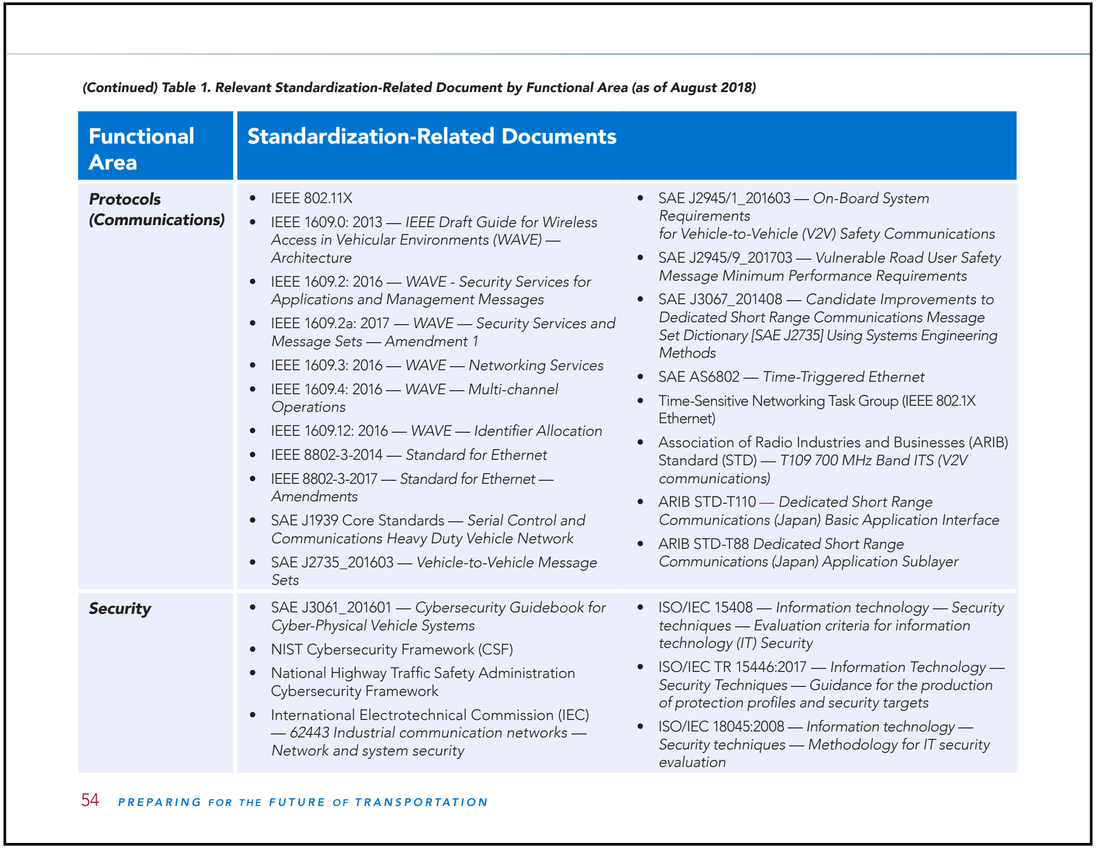


(Continued) Table 1. Relevant Standardization-Related Document by Functional Area (as of August 2018)

Functional Area	Standardization-Related Documents
Data	• Navigation Data Standard (NDS) — a standardized format for automotive-grade navigation databases, jointly developed by automobile manufacturers and suppliers. • North American Datum 1983 (NAD83) • World Geodetic System 1984 (WGS84) • European Terrestrial Reference System 1989 (ETRS89) • Chinese encrypted datum 2002 (CSJ-02) • ADASIS Forum vehicle to cloud messaging standards • Coordinated Universal Time (UTC) • International Atomic Time (TAI) • ISO 11270:2014 — *Intelligent Transport Systems — Lane Keeping Assistance Systems (LKAS) — Performance requirements and test procedures* • ISO 14296:2016 — *Intelligent Transport Systems — Extension of map database specifications for applications of cooperative Intelligent Transportation Systems* • ISO 14825:2011 — *Intelligent Transport Systems — Geographic Data Files (GDF) — GDF5.0* • ISO 15622:2010 — *Intelligent Transport Systems — Adaptive Cruise Control Systems — Performance requirements and test procedures* • ISO 19237:2017 — *Intelligent Transport Systems — Pedestrian detection and collision mitigation systems (PDCMS) — Performance requirements and test procedures* • ISO 22178:2009 — *Intelligent Transport Systems — Low speed following (LSF) systems — Performance requirements and test procedures* • ISO 22179:2009 — *Intelligent Transport Systems — Full Speed Range Adaptive (FSRA) systems — Performance requirements and test procedures* • ISO 22839:2013 — *Intelligent Transport Systems — Forward vehicle collision mitigation systems — Operation, performance, and verification requirements* • ISO/DIS 20035 — *Intelligent Transport Systems — Cooperative adaptive cruise control (CACC) — Operation, performance, and verification requirements* • SAE J1698 — *Event Data Recorder (EDR)*

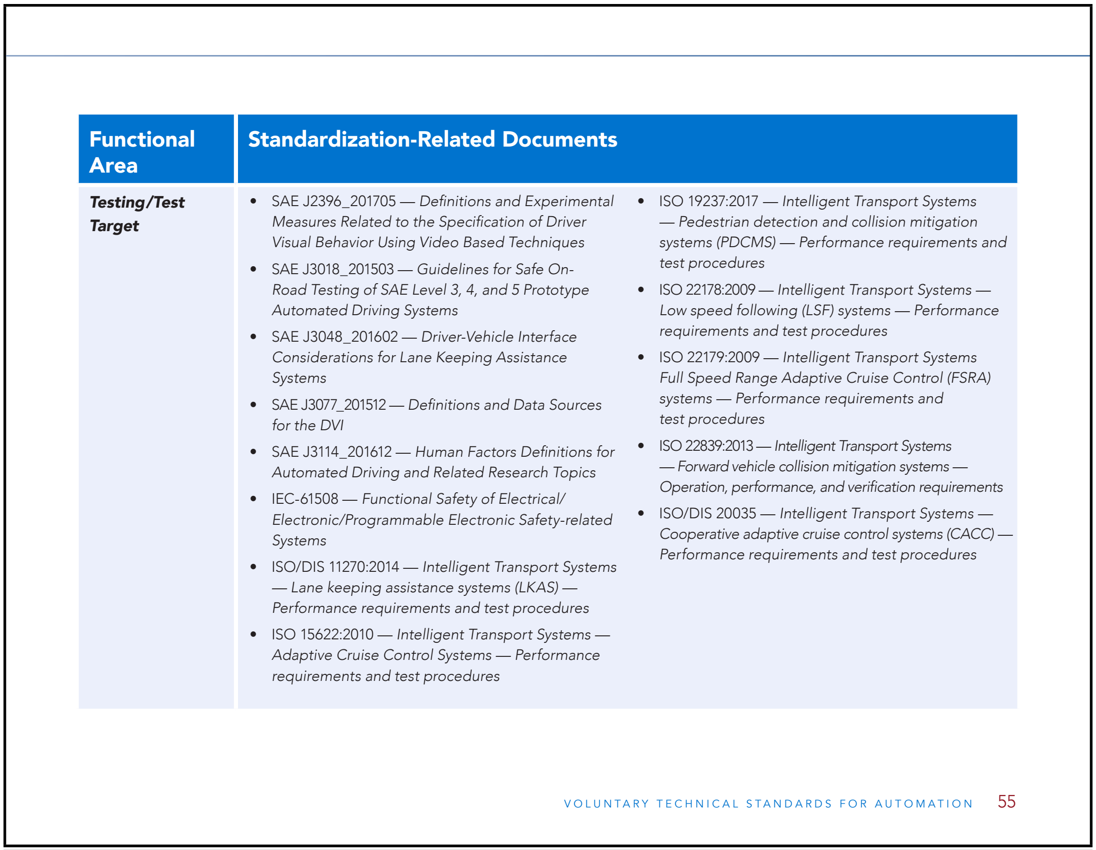


Functional Area	Standardization-Related Documents	
Design	- Federal Highway Administration *Manual on Uniform Traffic Control Devices (MUTCD)* - American Association of State Highway and Transportation Officials (AASHTO) *A Policy on Geometric Design of Highways and Streets* (Green Book) - *AASHTO Roadside Design Guide*	- Joint SAE-AASHTO *Committee on Road Markings* - ISO 2575:2010 — *Road vehicles — Symbols for controls, indicators, and tell-tales* - SAE J2945_201712 — *DSRC Systems Engineering Process Guidance for SAE J2945/X Documents and Common Design Concepts*
Maintenance and Inspections	- Commercial Vehicle Safety Alliance (CVSA) North American Standard Inspection Program (roadside inspection process for inspecting commercial motor vehicles and drivers throughout North America)	
Functional / Performance	- SAE J2958:2011 — *Report on Unmanned Ground Vehicle Reliability* - SAE J2980_201804 — *Considerations for ISO 26262 Automotive Safety Integrity Levels (ASIL) Hazard Classification* - SAE J3088 — *Active Safety System Sensors* - SAE J3116_201706 — *Active Safety Pedestrian Test Mannequin Recommendation* - U.S. Department of Defense (DOD) Military Standards (MIL-STD) — *882E Standard Practice for System Safety* - Radio Technical Commission for Aeronautics (RTCA) *DO-178C Software Considerations in Airborne Systems and Equipment Certification* - National Aeronautics and Space Administration (NASA) — *GB-8719.13 Software Safety Guidebook*	- Automated Driving and Platooning Task Force of the American Trucking Associations Technology and Maintenance Council - ISO 13482:2014 — *Robots and robotic devices — Safety requirements for personal care robots* - ISO 15622:2010 — *Intelligent Transport Systems — Adaptive Cruise Control systems — Performance requirements and test procedures* - ISO 17386:2010 — *Transport information and control systems — Maneuvering Aids for Low Speed Operation (MALSO) — Performance requirements and test procedures* - ISO 22840:2010 — *Intelligent Transport Systems — Devices to aid reverse maneuvers — Extended-range backing aid (ERBA) systems* - ISO 26262 — *Road vehicles — Functional safety*

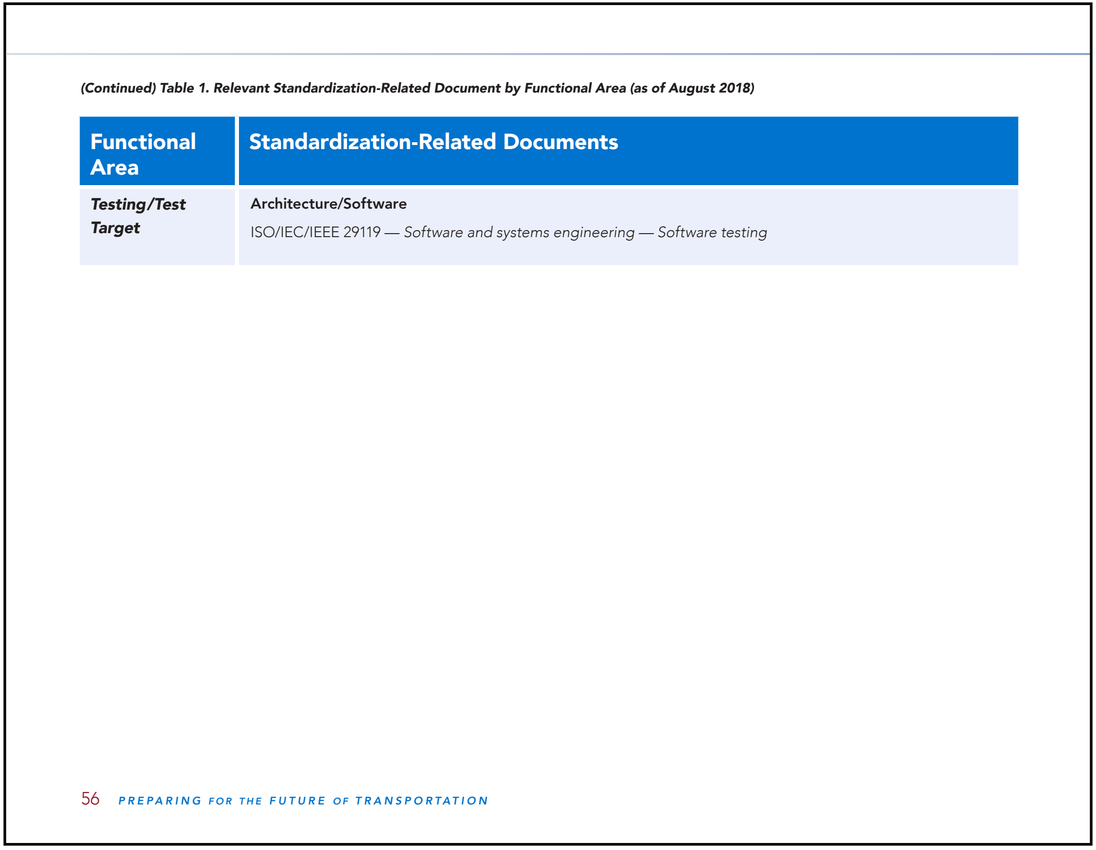


(Continued) Table 1. Relevant Standardization-Related Document by Functional Area (as of August 2018)

Functional Area	Standardization-Related Documents	
Protocols (Communications)	- IEEE 802.11X - IEEE 1609.0: 2013 — *IEEE Draft Guide for Wireless Access in Vehicular Environments (WAVE) — Architecture* - IEEE 1609.2: 2016 — *WAVE - Security Services for Applications and Management Messages* - IEEE 1609.2a: 2017 — *WAVE — Security Services and Message Sets — Amendment 1* - IEEE 1609.3: 2016 — *WAVE — Networking Services* - IEEE 1609.4: 2016 — *WAVE — Multi-channel Operations* - IEEE 1609.12: 2016 — *WAVE — Identifier Allocation* - IEEE 8802-3-2014 — *Standard for Ethernet* - IEEE 8802-3-2017 — *Standard for Ethernet — Amendments* - SAE J1939 Core Standards — *Serial Control and Communications Heavy Duty Vehicle Network* - SAE J2735_201603 — *Vehicle-to-Vehicle Message Sets*	- SAE J2945/1_201603 — *On-Board System Requirements for Vehicle-to-Vehicle (V2V) Safety Communications* - SAE J2945/9_201703 — *Vulnerable Road User Safety Message Minimum Performance Requirements* - SAE J3067_201408 — *Candidate Improvements to Dedicated Short Range Communications Message Set Dictionary [SAE J2735] Using Systems Engineering Methods* - SAE AS6802 — *Time-Triggered Ethernet* - Time-Sensitive Networking Task Group (IEEE 802.1X Ethernet) - Association of Radio Industries and Businesses (ARIB) Standard (STD) — *T109 700 MHz Band ITS (V2V communications)* - ARIB STD-T110 — *Dedicated Short Range Communications (Japan) Basic Application Interface* - ARIB STD-T88 *Dedicated Short Range Communications (Japan) Application Sublayer*
Security	- SAE J3061_201601 — *Cybersecurity Guidebook for Cyber-Physical Vehicle Systems* - NIST Cybersecurity Framework (CSF) - National Highway Traffic Safety Administration Cybersecurity Framework - International Electrotechnical Commission (IEC) — *62443 Industrial communication networks — Network and system security*	- ISO/IEC 15408 — *Information technology — Security techniques — Evaluation criteria for information technology (IT) Security* - ISO/IEC TR 15446:2017 — *Information Technology — Security Techniques — Guidance for the production of protection profiles and security targets* - ISO/IEC 18045:2008 — *Information technology — Security techniques — Methodology for IT security evaluation*

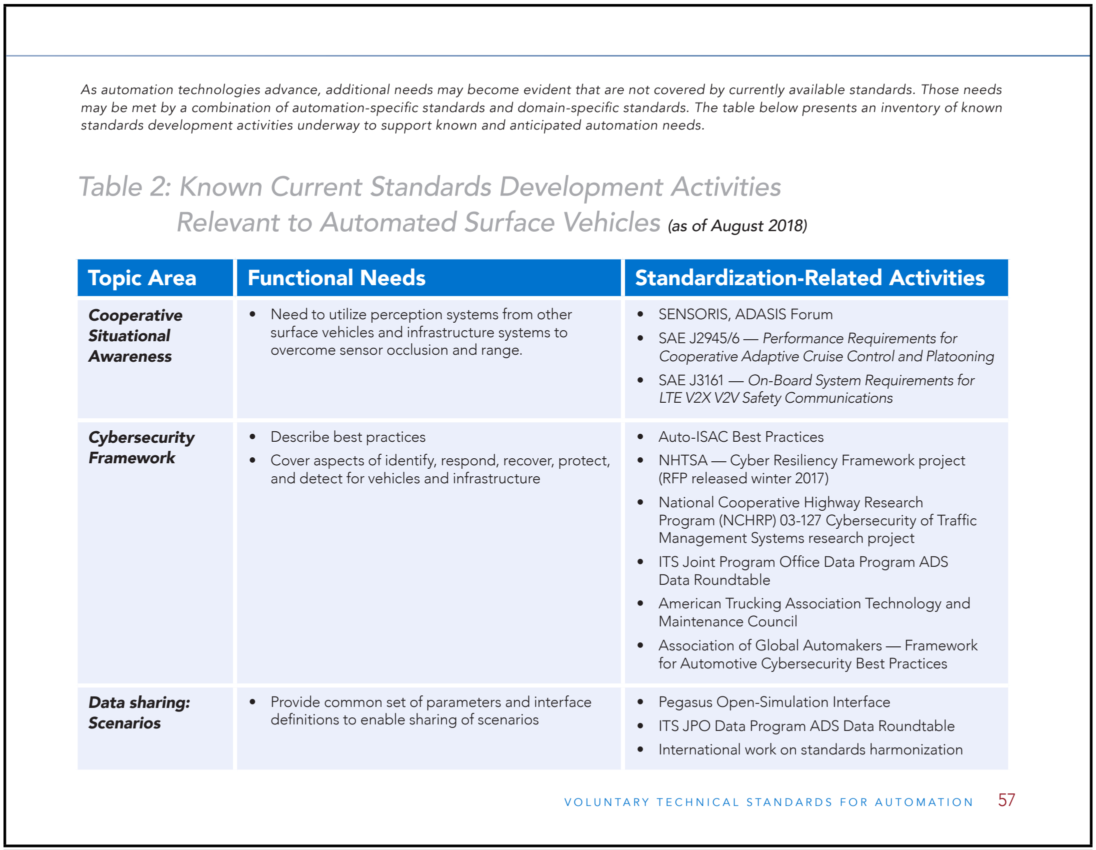


Functional Area	Standardization-Related Documents
Testing/Test Target	• SAE J2396_201705 — *Definitions and Experimental Measures Related to the Specification of Driver Visual Behavior Using Video Based Techniques* • SAE J3018_201503 — *Guidelines for Safe On-Road Testing of SAE Level 3, 4, and 5 Prototype Automated Driving Systems* • SAE J3048_201602 — *Driver-Vehicle Interface Considerations for Lane Keeping Assistance Systems* • SAE J3077_201512 — *Definitions and Data Sources for the DVI* • SAE J3114_201612 — *Human Factors Definitions for Automated Driving and Related Research Topics* • IEC-61508 — *Functional Safety of Electrical/Electronic/Programmable Electronic Safety-related Systems* • ISO/DIS 11270:2014 — *Intelligent Transport Systems — Lane keeping assistance systems (LKAS) — Performance requirements and test procedures* • ISO 15622:2010 — *Intelligent Transport Systems — Adaptive Cruise Control Systems — Performance requirements and test procedures* • ISO 19237:2017 — *Intelligent Transport Systems — Pedestrian detection and collision mitigation systems (PDCMS) — Performance requirements and test procedures* • ISO 22178:2009 — *Intelligent Transport Systems — Low speed following (LSF) systems — Performance requirements and test procedures* • ISO 22179:2009 — *Intelligent Transport Systems Full Speed Range Adaptive Cruise Control (FSRA) systems — Performance requirements and test procedures* • ISO 22839:2013 — *Intelligent Transport Systems — Forward vehicle collision mitigation systems — Operation, performance, and verification requirements* • ISO/DIS 20035 — *Intelligent Transport Systems — Cooperative adaptive cruise control systems (CACC) — Performance requirements and test procedures*



(Continued) Table 1. Relevant Standardization-Related Document by Functional Area (as of August 2018)

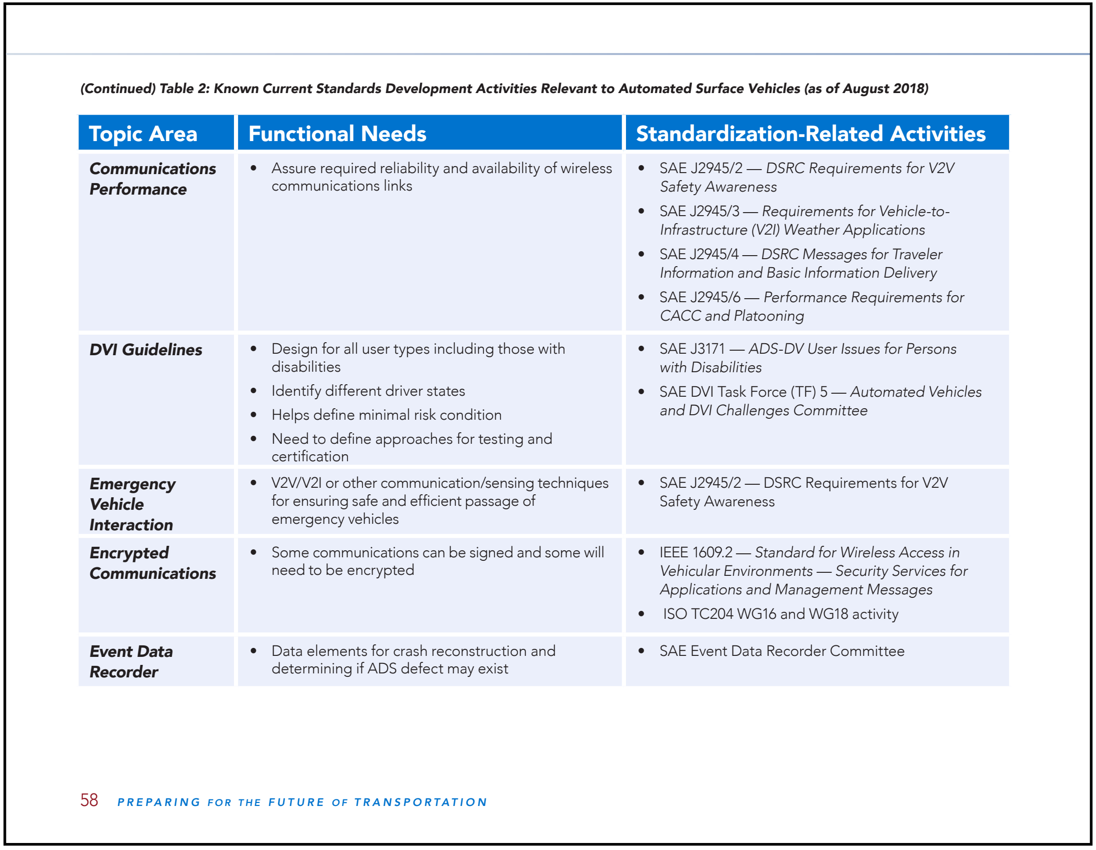
Functional Area	Standardization-Related Documents
Testing/Test Target	Architecture/Software ISO/IEC/IEEE 29119 — *Software and systems engineering — Software testing*



As automation technologies advance, additional needs may become evident that are not covered by currently available standards. Those needs may be met by a combination of automation-specific standards and domain-specific standards. The table below presents an inventory of known standards development activities underway to support known and anticipated automation needs.

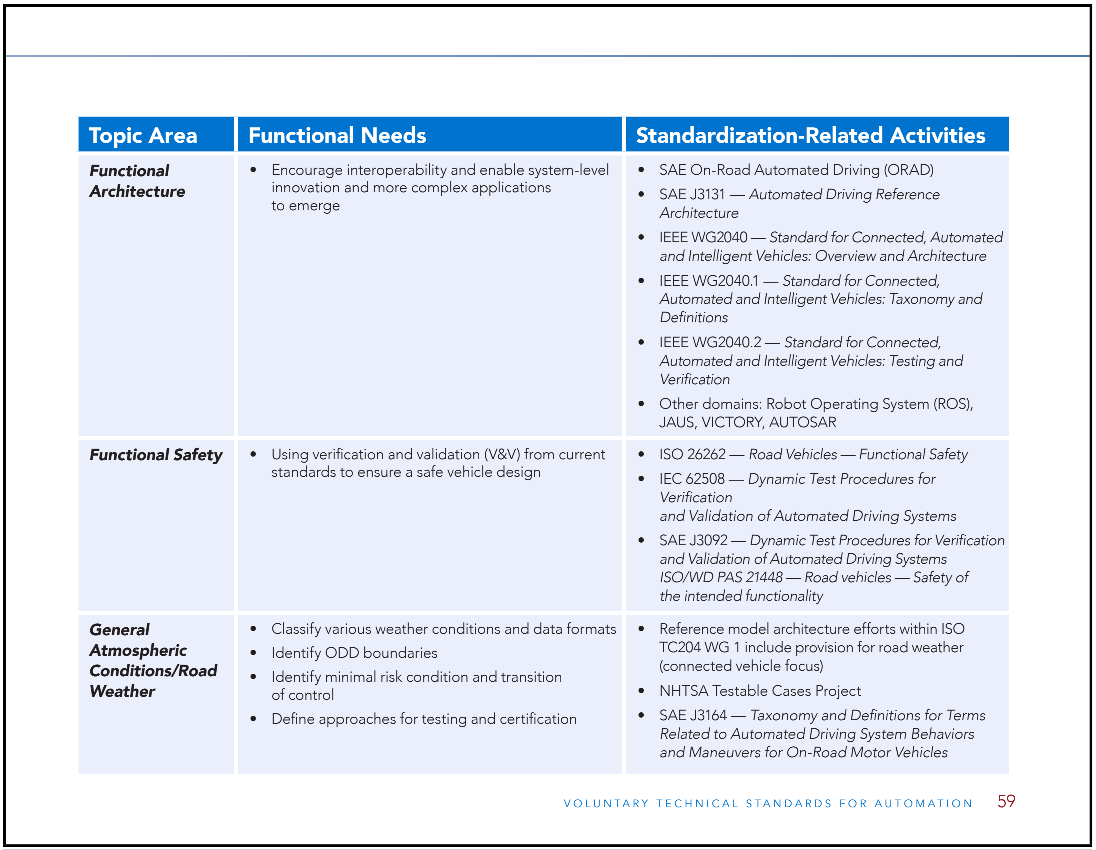
Table 2: Known Current Standards Development Activities Relevant to Automated Surface Vehicles *(as of August 2018)*

Topic Area	Functional Needs	Standardization-Related Activities
Cooperative Situational Awareness	• Need to utilize perception systems from other surface vehicles and infrastructure systems to overcome sensor occlusion and range.	• SENSORIS, ADASIS Forum • SAE J2945/6 — *Performance Requirements for Cooperative Adaptive Cruise Control and Platooning* • SAE J3161 — *On-Board System Requirements for LTE V2X V2V Safety Communications*
Cybersecurity Framework	• Describe best practices • Cover aspects of identify, respond, recover, protect, and detect for vehicles and infrastructure	• Auto-ISAC Best Practices • NHTSA — Cyber Resiliency Framework project (RFP released winter 2017) • National Cooperative Highway Research Program (NCHRP) 03-127 Cybersecurity of Traffic Management Systems research project • ITS Joint Program Office Data Program ADS Data Roundtable • American Trucking Association Technology and Maintenance Council • Association of Global Automakers — Framework for Automotive Cybersecurity Best Practices
Data sharing: Scenarios	• Provide common set of parameters and interface definitions to enable sharing of scenarios	• Pegasus Open-Simulation Interface • ITS JPO Data Program ADS Data Roundtable • International work on standards harmonization

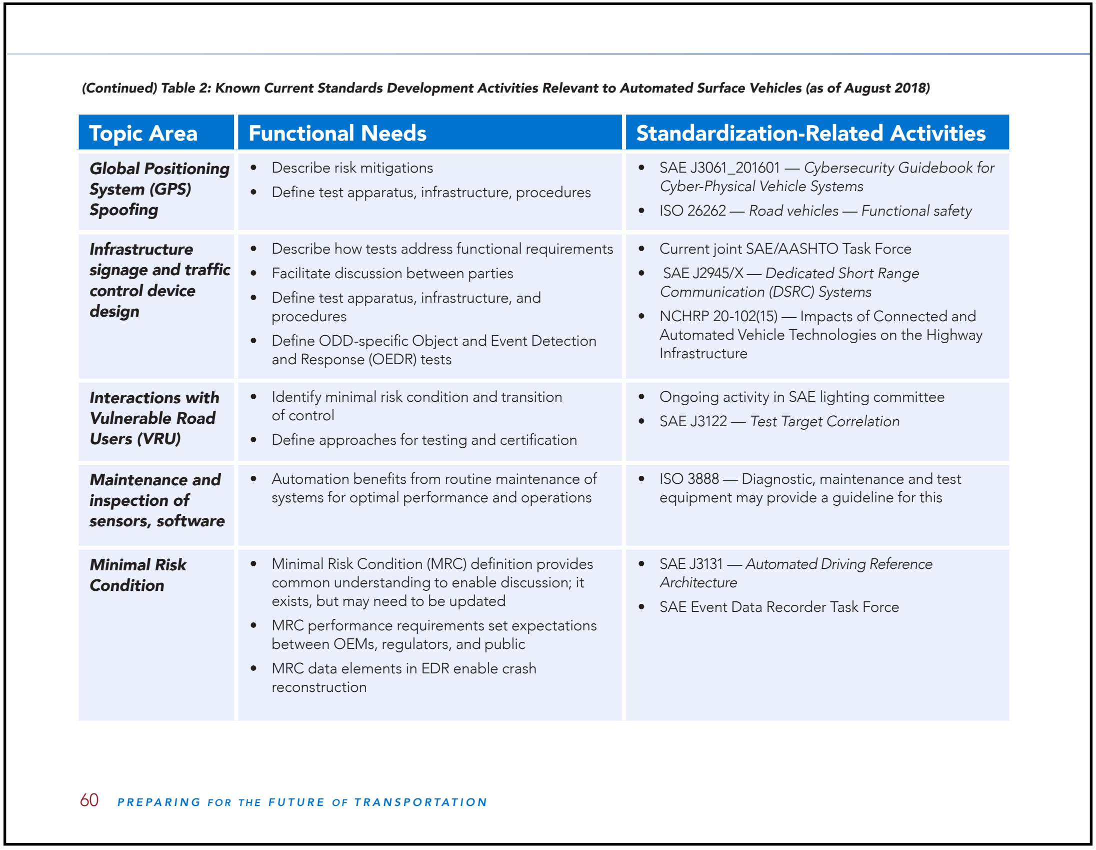


(Continued) Table 2: Known Current Standards Development Activities Relevant to Automated Surface Vehicles (as of August 2018)

Topic Area	Functional Needs	Standardization-Related Activities
Communications Performance	• Assure required reliability and availability of wireless communications links	• SAE J2945/2 — *DSRC Requirements for V2V Safety Awareness* • SAE J2945/3 — *Requirements for Vehicle-to-Infrastructure (V2I) Weather Applications* • SAE J2945/4 — *DSRC Messages for Traveler Information and Basic Information Delivery* • SAE J2945/6 — *Performance Requirements for CACC and Platooning*
DVI Guidelines	• Design for all user types including those with disabilities • Identify different driver states • Helps define minimal risk condition • Need to define approaches for testing and certification	• SAE J3171 — *ADS-DV User Issues for Persons with Disabilities* • SAE DVI Task Force (TF) 5 — *Automated Vehicles and DVI Challenges Committee*
Emergency Vehicle Interaction	• V2V/V2I or other communication/sensing techniques for ensuring safe and efficient passage of emergency vehicles	• SAE J2945/2 — DSRC Requirements for V2V Safety Awareness
Encrypted Communications	• Some communications can be signed and some will need to be encrypted	• IEEE 1609.2 — *Standard for Wireless Access in Vehicular Environments — Security Services for Applications and Management Messages* • ISO TC204 WG16 and WG18 activity
Event Data Recorder	• Data elements for crash reconstruction and determining if ADS defect may exist	• SAE Event Data Recorder Committee

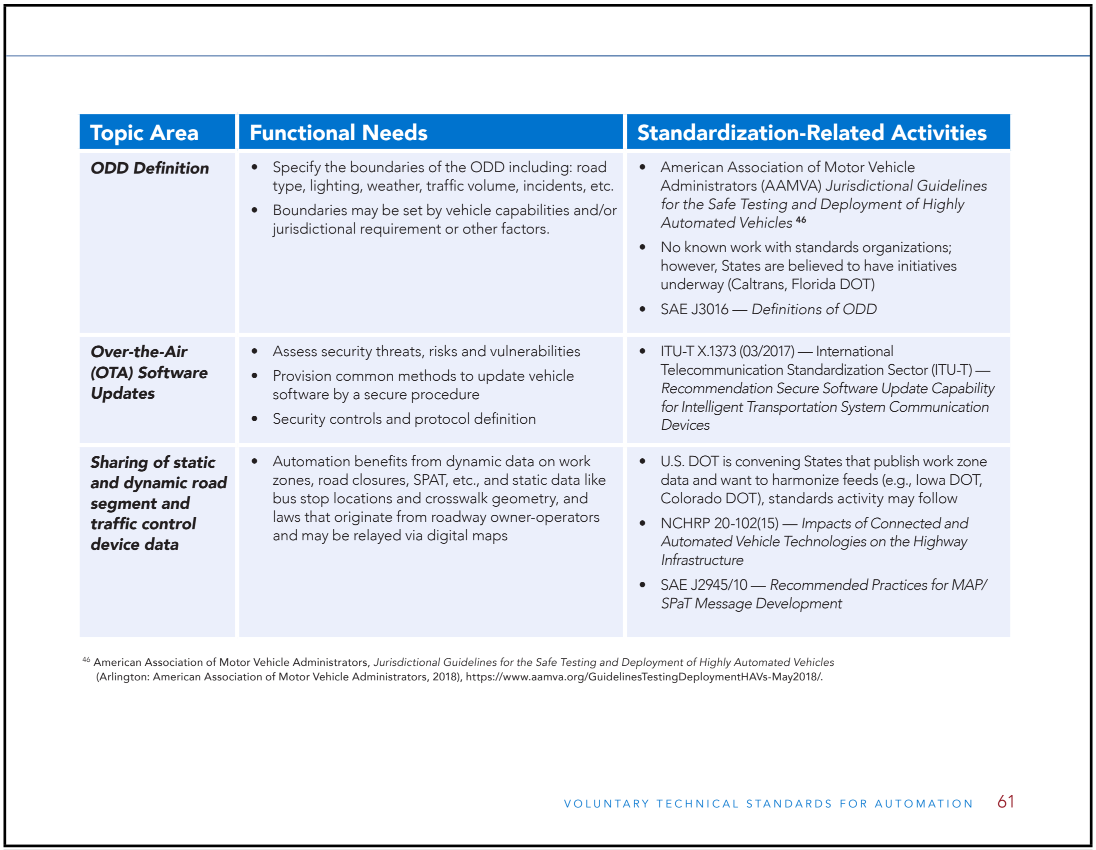


Topic Area	Functional Needs	Standardization-Related Activities
Functional Architecture	• Encourage interoperability and enable system-level innovation and more complex applications to emerge	• SAE On-Road Automated Driving (ORAD) • SAE J3131 — *Automated Driving Reference Architecture* • IEEE WG2040 — *Standard for Connected, Automated and Intelligent Vehicles: Overview and Architecture* • IEEE WG2040.1 — *Standard for Connected, Automated and Intelligent Vehicles: Taxonomy and Definitions* • IEEE WG2040.2 — *Standard for Connected, Automated and Intelligent Vehicles: Testing and Verification* • Other domains: Robot Operating System (ROS), JAUS, VICTORY, AUTOSAR
Functional Safety	• Using verification and validation (V&V) from current standards to ensure a safe vehicle design	• ISO 26262 — *Road Vehicles — Functional Safety* • IEC 62508 — *Dynamic Test Procedures for Verification and Validation of Automated Driving Systems* • SAE J3092 — *Dynamic Test Procedures for Verification and Validation of Automated Driving Systems* *ISO/WD PAS 21448 — Road vehicles — Safety of the intended functionality*
General Atmospheric Conditions/Road Weather	• Classify various weather conditions and data formats • Identify ODD boundaries • Identify minimal risk condition and transition of control • Define approaches for testing and certification	• Reference model architecture efforts within ISO TC204 WG 1 include provision for road weather (connected vehicle focus) • NHTSA Testable Cases Project • SAE J3164 — *Taxonomy and Definitions for Terms Related to Automated Driving System Behaviors and Maneuvers for On-Road Motor Vehicles*



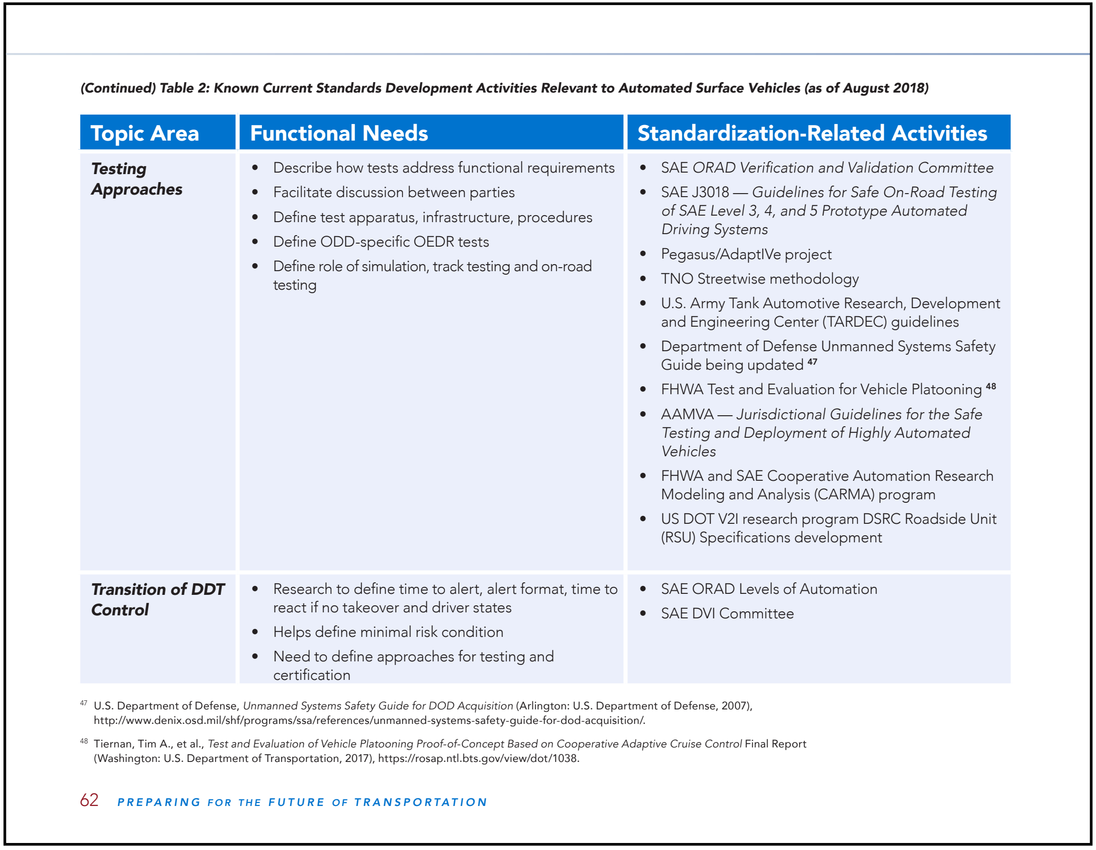
(Continued) Table 2: Known Current Standards Development Activities Relevant to Automated Surface Vehicles (as of August 2018)

Topic Area	Functional Needs	Standardization-Related Activities
Global Positioning System (GPS) Spoofing	• Describe risk mitigations • Define test apparatus, infrastructure, procedures	• SAE J3061_201601 — *Cybersecurity Guidebook for Cyber-Physical Vehicle Systems* • ISO 26262 — *Road vehicles — Functional safety*
Infrastructure signage and traffic control device design	• Describe how tests address functional requirements • Facilitate discussion between parties • Define test apparatus, infrastructure, and procedures • Define ODD-specific Object and Event Detection and Response (OEDR) tests	• Current joint SAE/AASHTO Task Force • SAE J2945/X — *Dedicated Short Range Communication (DSRC) Systems* • NCHRP 20-102(15) — Impacts of Connected and Automated Vehicle Technologies on the Highway Infrastructure
Interactions with Vulnerable Road Users (VRU)	• Identify minimal risk condition and transition of control • Define approaches for testing and certification	• Ongoing activity in SAE lighting committee • SAE J3122 — *Test Target Correlation*
Maintenance and inspection of sensors, software	• Automation benefits from routine maintenance of systems for optimal performance and operations	• ISO 3888 — Diagnostic, maintenance and test equipment may provide a guideline for this
Minimal Risk Condition	• Minimal Risk Condition (MRC) definition provides common understanding to enable discussion; it exists, but may need to be updated • MRC performance requirements set expectations between OEMs, regulators, and public • MRC data elements in EDR enable crash reconstruction	• SAE J3131 — *Automated Driving Reference Architecture* • SAE Event Data Recorder Task Force



Topic Area	Functional Needs	Standardization-Related Activities
ODD Definition	• Specify the boundaries of the ODD including: road type, lighting, weather, traffic volume, incidents, etc. • Boundaries may be set by vehicle capabilities and/or jurisdictional requirement or other factors.	• American Association of Motor Vehicle Administrators (AAMVA) *Jurisdictional Guidelines for the Safe Testing and Deployment of Highly Automated Vehicles* [46] • No known work with standards organizations; however, States are believed to have initiatives underway (Caltrans, Florida DOT) • SAE J3016 — *Definitions of ODD*
Over-the-Air (OTA) Software Updates	• Assess security threats, risks and vulnerabilities • Provision common methods to update vehicle software by a secure procedure • Security controls and protocol definition	• ITU-T X.1373 (03/2017) — International Telecommunication Standardization Sector (ITU-T) — *Recommendation Secure Software Update Capability for Intelligent Transportation System Communication Devices*
Sharing of static and dynamic road segment and traffic control device data	• Automation benefits from dynamic data on work zones, road closures, SPAT, etc., and static data like bus stop locations and crosswalk geometry, and laws that originate from roadway owner-operators and may be relayed via digital maps	• U.S. DOT is convening States that publish work zone data and want to harmonize feeds (e.g., Iowa DOT, Colorado DOT), standards activity may follow • NCHRP 20-102(15) — *Impacts of Connected and Automated Vehicle Technologies on the Highway Infrastructure* • SAE J2945/10 — *Recommended Practices for MAP/SPaT Message Development*

[46] American Association of Motor Vehicle Administrators, *Jurisdictional Guidelines for the Safe Testing and Deployment of Highly Automated Vehicles* (Arlington: American Association of Motor Vehicle Administrators, 2018), https://www.aamva.org/GuidelinesTestingDeploymentHAVs-May2018/.

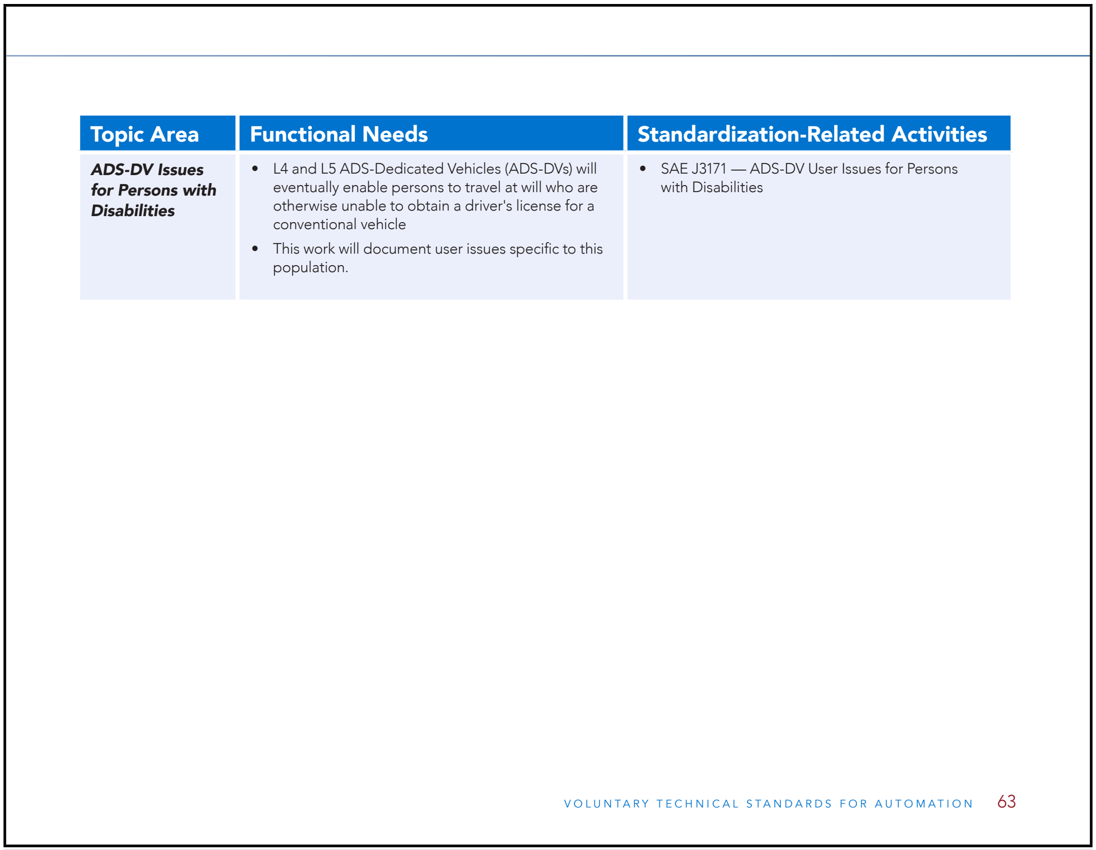


(Continued) Table 2: Known Current Standards Development Activities Relevant to Automated Surface Vehicles (as of August 2018)

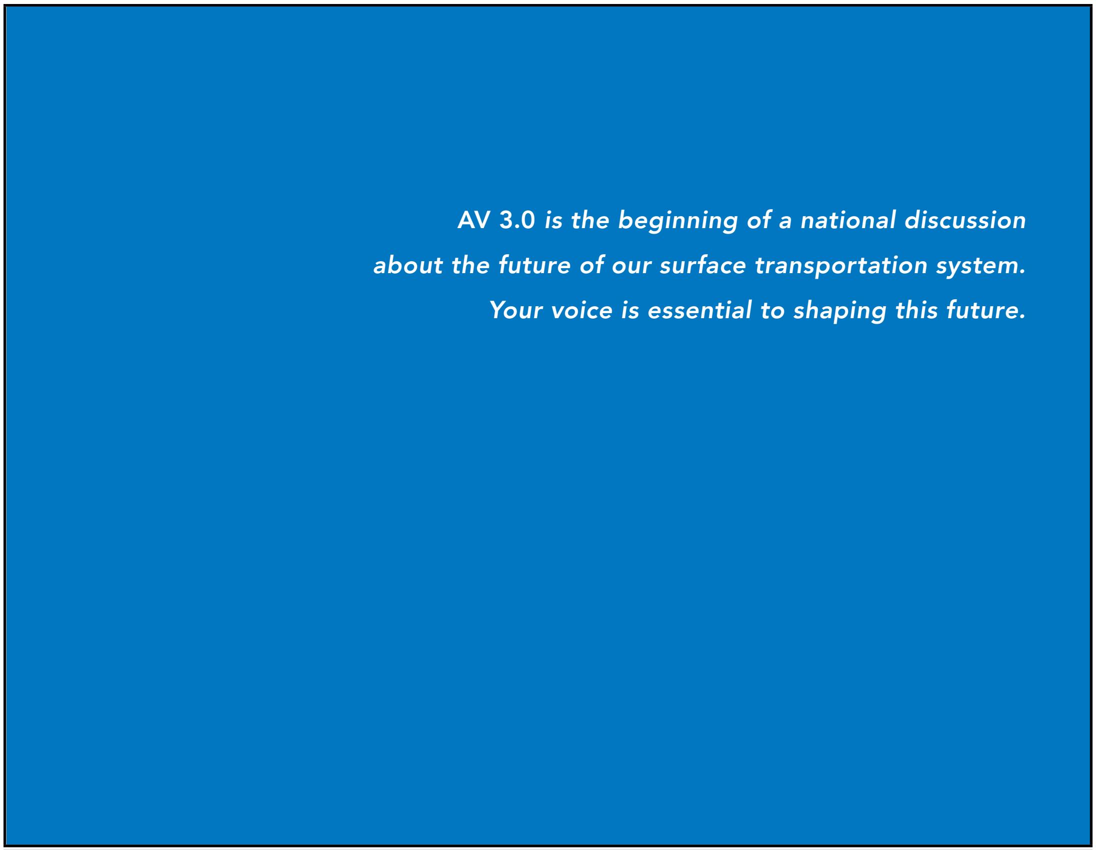
Topic Area	Functional Needs	Standardization-Related Activities
Testing Approaches	• Describe how tests address functional requirements • Facilitate discussion between parties • Define test apparatus, infrastructure, procedures • Define ODD-specific OEDR tests • Define role of simulation, track testing and on-road testing	• SAE *ORAD Verification and Validation Committee* • SAE J3018 — *Guidelines for Safe On-Road Testing of SAE Level 3, 4, and 5 Prototype Automated Driving Systems* • Pegasus/AdaptIVe project • TNO Streetwise methodology • U.S. Army Tank Automotive Research, Development and Engineering Center (TARDEC) guidelines • Department of Defense Unmanned Systems Safety Guide being updated [47] • FHWA Test and Evaluation for Vehicle Platooning [48] • AAMVA — *Jurisdictional Guidelines for the Safe Testing and Deployment of Highly Automated Vehicles* • FHWA and SAE Cooperative Automation Research Modeling and Analysis (CARMA) program • US DOT V2I research program DSRC Roadside Unit (RSU) Specifications development
Transition of DDT Control	• Research to define time to alert, alert format, time to react if no takeover and driver states • Helps define minimal risk condition • Need to define approaches for testing and certification	• SAE ORAD Levels of Automation • SAE DVI Committee

[47] U.S. Department of Defense, *Unmanned Systems Safety Guide for DOD Acquisition* (Arlington: U.S. Department of Defense, 2007), http://www.denix.osd.mil/shf/programs/ssa/references/unmanned-systems-safety-guide-for-dod-acquisition/.

[48] Tiernan, Tim A., et al., *Test and Evaluation of Vehicle Platooning Proof-of-Concept Based on Cooperative Adaptive Cruise Control* Final Report (Washington: U.S. Department of Transportation, 2017), https://rosap.ntl.bts.gov/view/dot/1038.



Topic Area	Functional Needs	Standardization-Related Activities
ADS-DV Issues for Persons with Disabilities	• L4 and L5 ADS-Dedicated Vehicles (ADS-DVs) will eventually enable persons to travel at will who are otherwise unable to obtain a driver's license for a conventional vehicle • This work will document user issues specific to this population.	• SAE J3171 — ADS-DV User Issues for Persons with Disabilities



AV 3.0 is the beginning of a national discussion about the future of our surface transportation system. Your voice is essential to shaping this future.

Preparing for

THE FUTURE OF TRANSPORTATION

Automated Vehicles 3.0

October 2018
https://www.transportation.gov/av

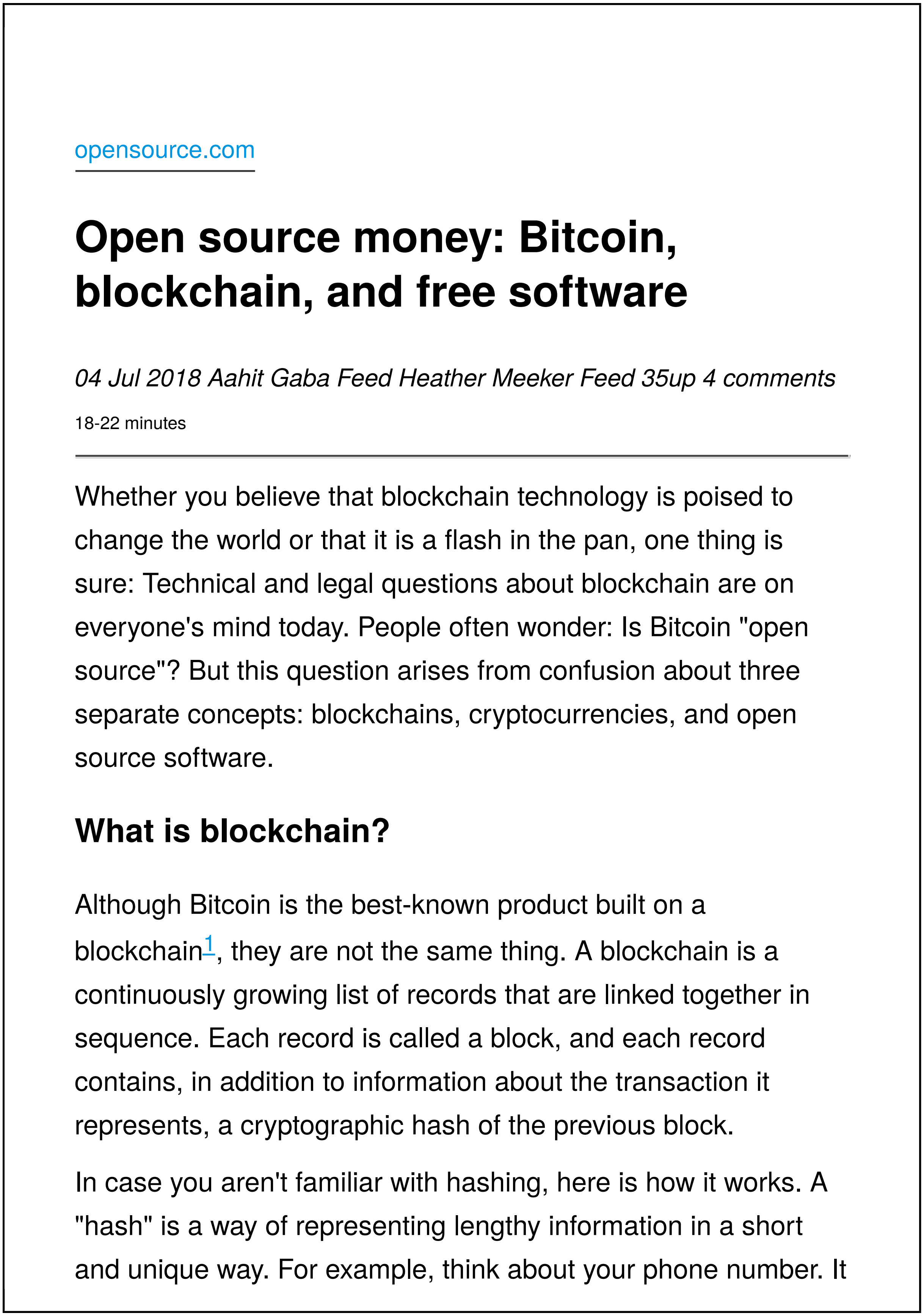
U.S. Department of Transportation

Part IV

FOSS, Blockchain and AI

Chapter 18

Open source money: Bitcoin, blockchain, and free software (Meeker and Gaba)



opensource.com

Open source money: Bitcoin, blockchain, and free software

04 Jul 2018 Aahit Gaba Feed Heather Meeker Feed 35up 4 comments

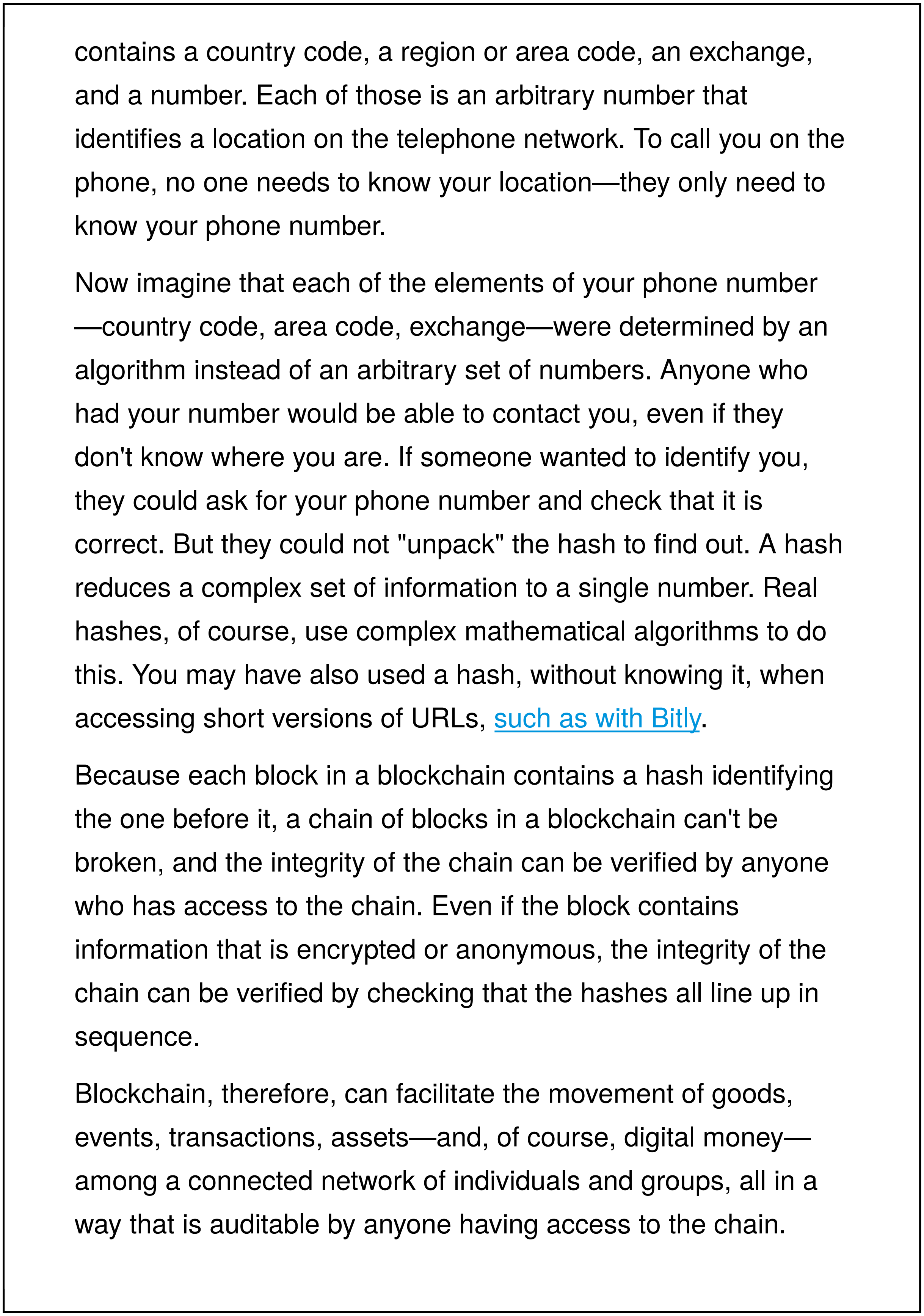
18-22 minutes

Whether you believe that blockchain technology is poised to change the world or that it is a flash in the pan, one thing is sure: Technical and legal questions about blockchain are on everyone's mind today. People often wonder: Is Bitcoin "open source"? But this question arises from confusion about three separate concepts: blockchains, cryptocurrencies, and open source software.

What is blockchain?

Although Bitcoin is the best-known product built on a blockchain[1], they are not the same thing. A blockchain is a continuously growing list of records that are linked together in sequence. Each record is called a block, and each record contains, in addition to information about the transaction it represents, a cryptographic hash of the previous block.

In case you aren't familiar with hashing, here is how it works. A "hash" is a way of representing lengthy information in a short and unique way. For example, think about your phone number. It

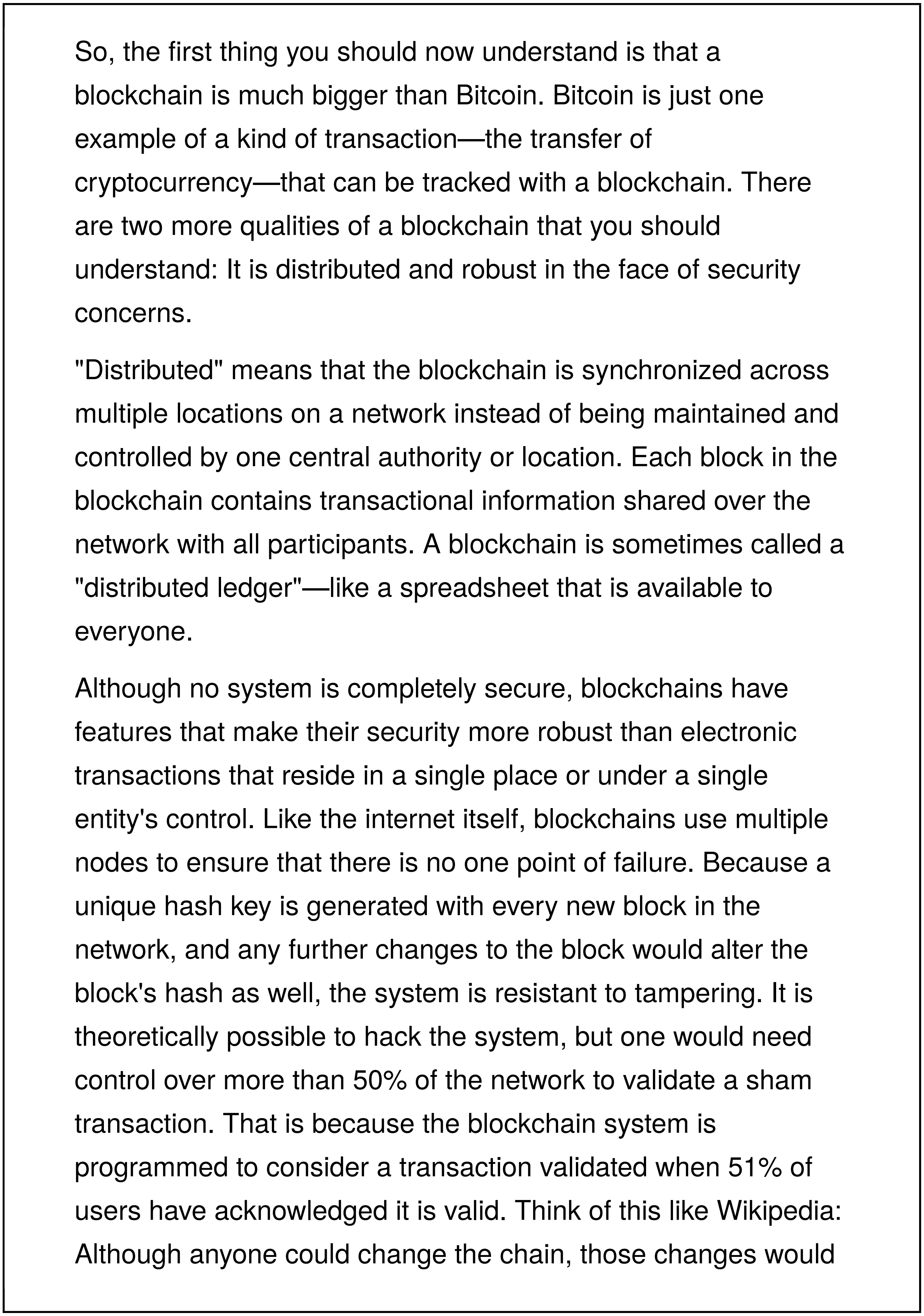


contains a country code, a region or area code, an exchange, and a number. Each of those is an arbitrary number that identifies a location on the telephone network. To call you on the phone, no one needs to know your location—they only need to know your phone number.

Now imagine that each of the elements of your phone number —country code, area code, exchange—were determined by an algorithm instead of an arbitrary set of numbers. Anyone who had your number would be able to contact you, even if they don't know where you are. If someone wanted to identify you, they could ask for your phone number and check that it is correct. But they could not "unpack" the hash to find out. A hash reduces a complex set of information to a single number. Real hashes, of course, use complex mathematical algorithms to do this. You may have also used a hash, without knowing it, when accessing short versions of URLs, such as with Bitly.

Because each block in a blockchain contains a hash identifying the one before it, a chain of blocks in a blockchain can't be broken, and the integrity of the chain can be verified by anyone who has access to the chain. Even if the block contains information that is encrypted or anonymous, the integrity of the chain can be verified by checking that the hashes all line up in sequence.

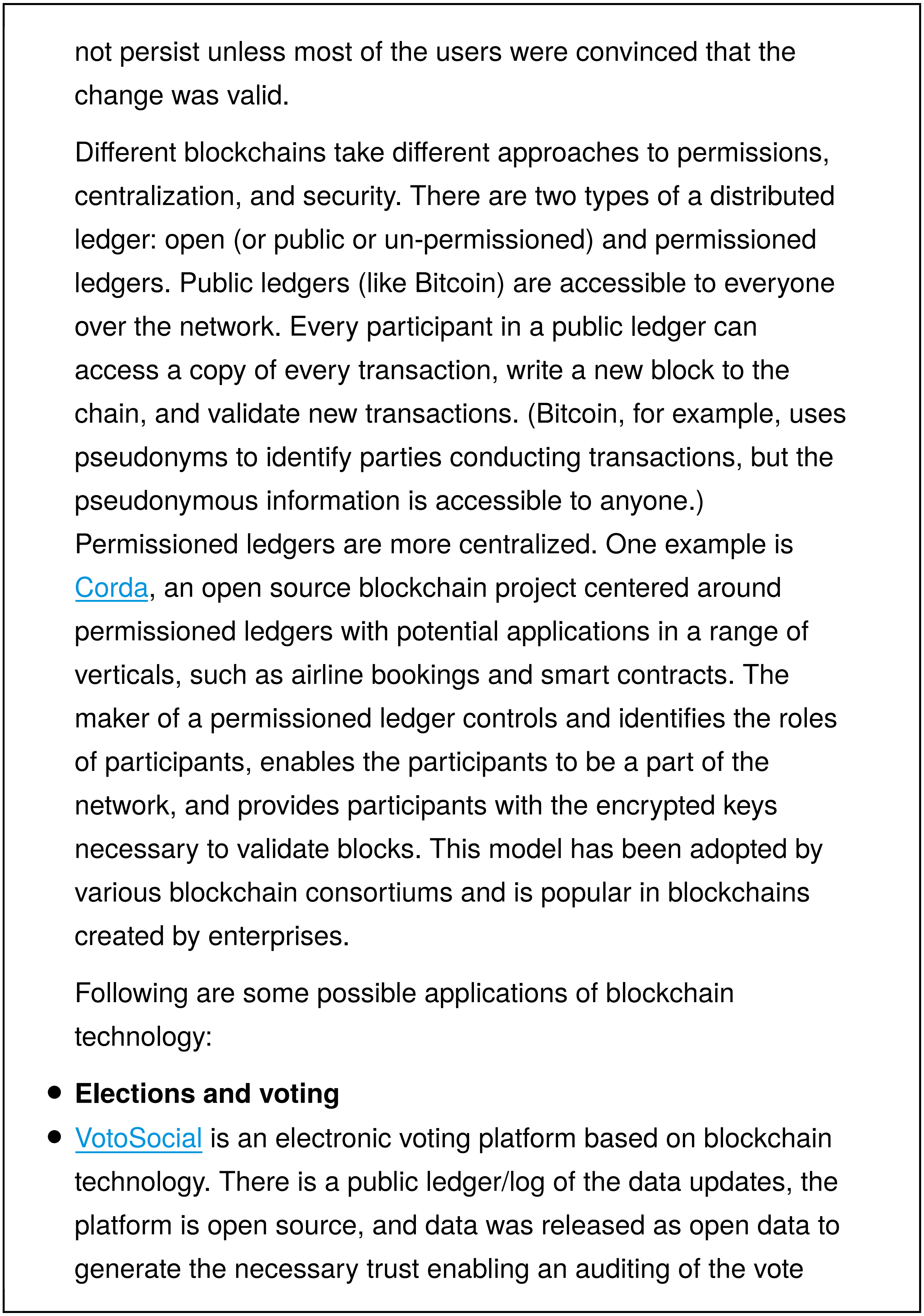
Blockchain, therefore, can facilitate the movement of goods, events, transactions, assets—and, of course, digital money— among a connected network of individuals and groups, all in a way that is auditable by anyone having access to the chain.



So, the first thing you should now understand is that a blockchain is much bigger than Bitcoin. Bitcoin is just one example of a kind of transaction—the transfer of cryptocurrency—that can be tracked with a blockchain. There are two more qualities of a blockchain that you should understand: It is distributed and robust in the face of security concerns.

"Distributed" means that the blockchain is synchronized across multiple locations on a network instead of being maintained and controlled by one central authority or location. Each block in the blockchain contains transactional information shared over the network with all participants. A blockchain is sometimes called a "distributed ledger"—like a spreadsheet that is available to everyone.

Although no system is completely secure, blockchains have features that make their security more robust than electronic transactions that reside in a single place or under a single entity's control. Like the internet itself, blockchains use multiple nodes to ensure that there is no one point of failure. Because a unique hash key is generated with every new block in the network, and any further changes to the block would alter the block's hash as well, the system is resistant to tampering. It is theoretically possible to hack the system, but one would need control over more than 50% of the network to validate a sham transaction. That is because the blockchain system is programmed to consider a transaction validated when 51% of users have acknowledged it is valid. Think of this like Wikipedia: Although anyone could change the chain, those changes would

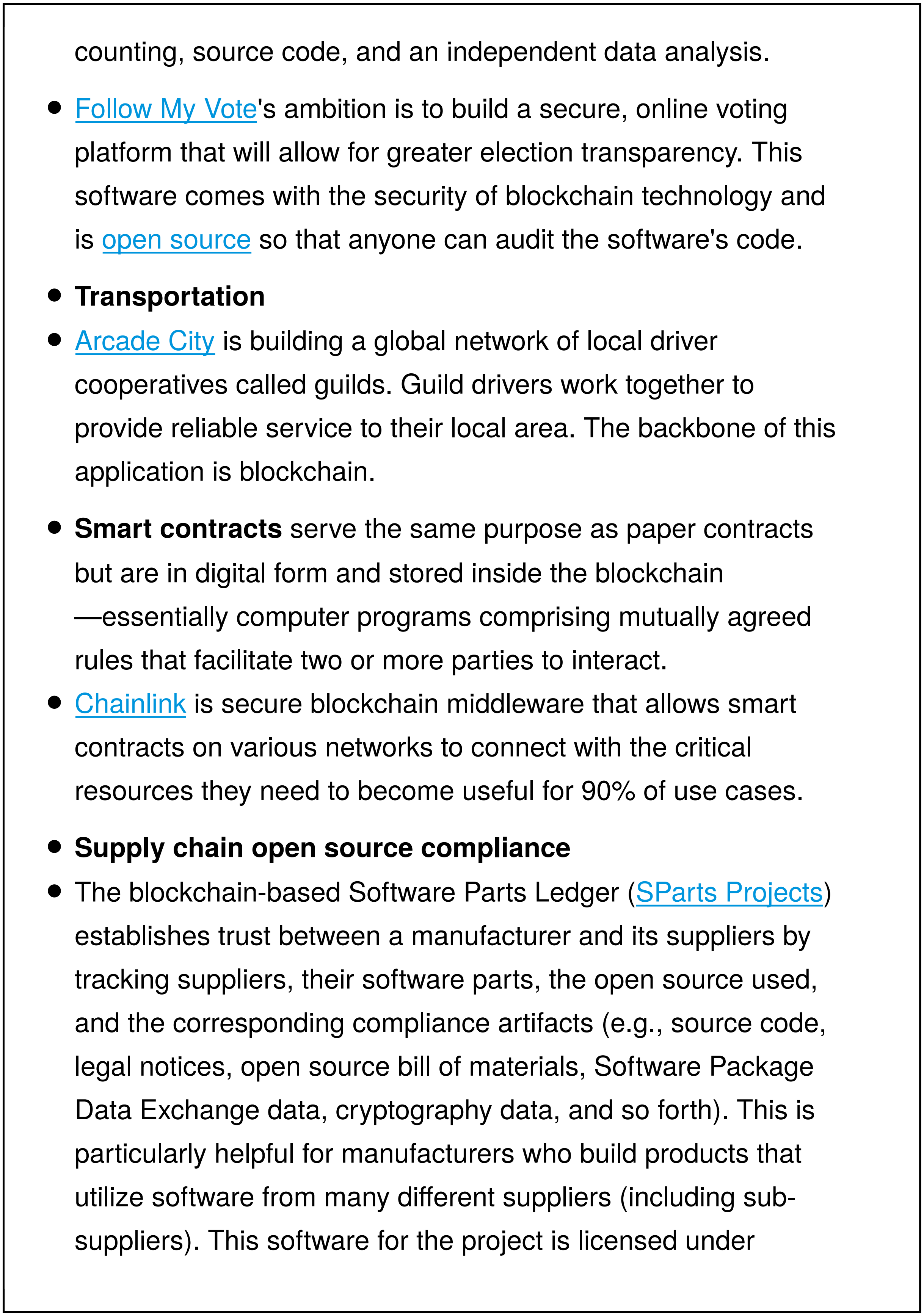


not persist unless most of the users were convinced that the change was valid.

Different blockchains take different approaches to permissions, centralization, and security. There are two types of a distributed ledger: open (or public or un-permissioned) and permissioned ledgers. Public ledgers (like Bitcoin) are accessible to everyone over the network. Every participant in a public ledger can access a copy of every transaction, write a new block to the chain, and validate new transactions. (Bitcoin, for example, uses pseudonyms to identify parties conducting transactions, but the pseudonymous information is accessible to anyone.) Permissioned ledgers are more centralized. One example is Corda, an open source blockchain project centered around permissioned ledgers with potential applications in a range of verticals, such as airline bookings and smart contracts. The maker of a permissioned ledger controls and identifies the roles of participants, enables the participants to be a part of the network, and provides participants with the encrypted keys necessary to validate blocks. This model has been adopted by various blockchain consortiums and is popular in blockchains created by enterprises.

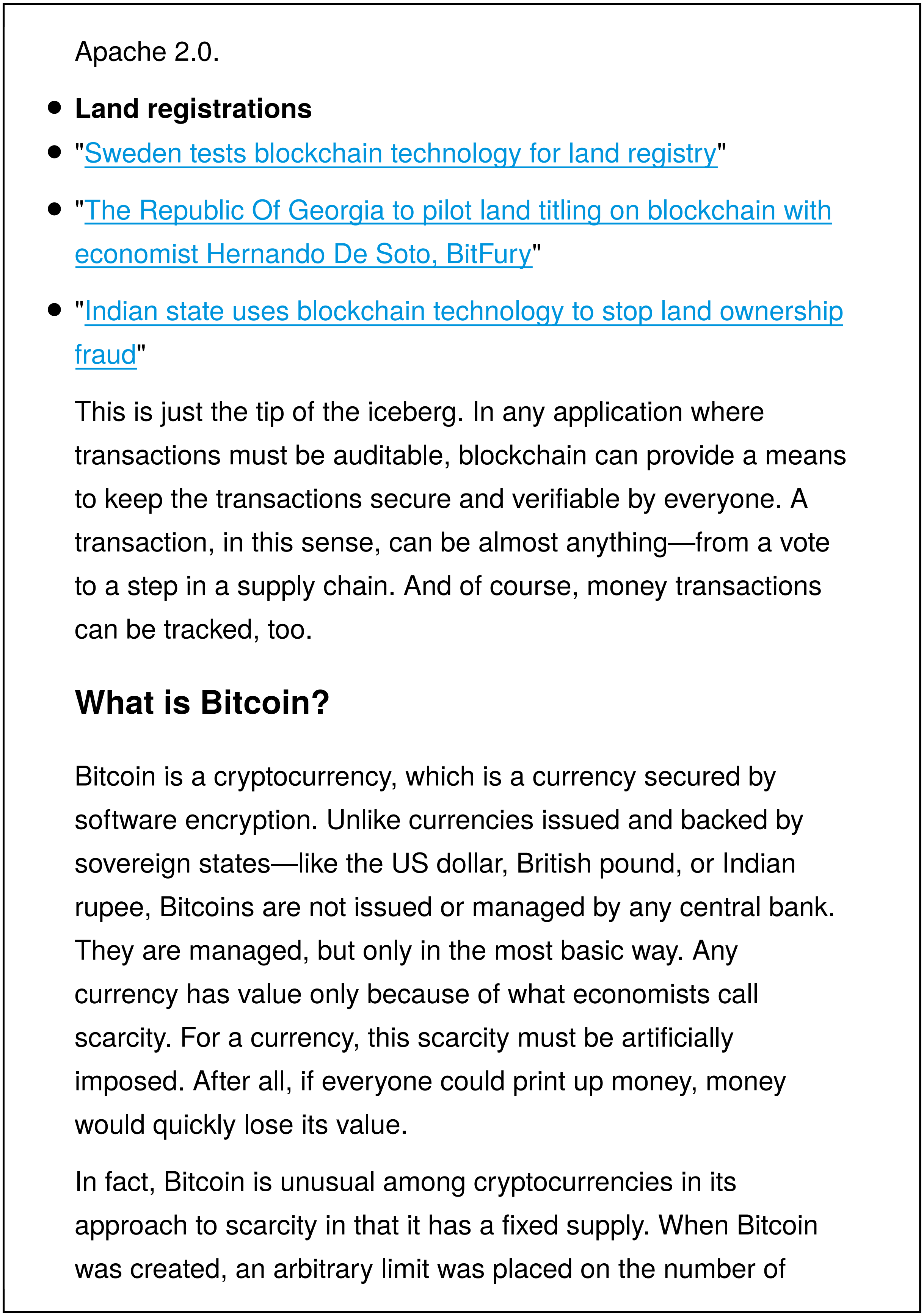
Following are some possible applications of blockchain technology:

- **Elections and voting**
- VotoSocial is an electronic voting platform based on blockchain technology. There is a public ledger/log of the data updates, the platform is open source, and data was released as open data to generate the necessary trust enabling an auditing of the vote



counting, source code, and an independent data analysis.

- Follow My Vote's ambition is to build a secure, online voting platform that will allow for greater election transparency. This software comes with the security of blockchain technology and is open source so that anyone can audit the software's code.
- **Transportation**
- Arcade City is building a global network of local driver cooperatives called guilds. Guild drivers work together to provide reliable service to their local area. The backbone of this application is blockchain.
- **Smart contracts** serve the same purpose as paper contracts but are in digital form and stored inside the blockchain —essentially computer programs comprising mutually agreed rules that facilitate two or more parties to interact.
- Chainlink is secure blockchain middleware that allows smart contracts on various networks to connect with the critical resources they need to become useful for 90% of use cases.
- **Supply chain open source compliance**
- The blockchain-based Software Parts Ledger (SParts Projects) establishes trust between a manufacturer and its suppliers by tracking suppliers, their software parts, the open source used, and the corresponding compliance artifacts (e.g., source code, legal notices, open source bill of materials, Software Package Data Exchange data, cryptography data, and so forth). This is particularly helpful for manufacturers who build products that utilize software from many different suppliers (including sub-suppliers). This software for the project is licensed under



Apache 2.0.

- **Land registrations**
- "Sweden tests blockchain technology for land registry"
- "The Republic Of Georgia to pilot land titling on blockchain with economist Hernando De Soto, BitFury"
- "Indian state uses blockchain technology to stop land ownership fraud"

This is just the tip of the iceberg. In any application where transactions must be auditable, blockchain can provide a means to keep the transactions secure and verifiable by everyone. A transaction, in this sense, can be almost anything—from a vote to a step in a supply chain. And of course, money transactions can be tracked, too.

What is Bitcoin?

Bitcoin is a cryptocurrency, which is a currency secured by software encryption. Unlike currencies issued and backed by sovereign states—like the US dollar, British pound, or Indian rupee, Bitcoins are not issued or managed by any central bank. They are managed, but only in the most basic way. Any currency has value only because of what economists call scarcity. For a currency, this scarcity must be artificially imposed. After all, if everyone could print up money, money would quickly lose its value.

In fact, Bitcoin is unusual among cryptocurrencies in its approach to scarcity in that it has a fixed supply. When Bitcoin was created, an arbitrary limit was placed on the number of

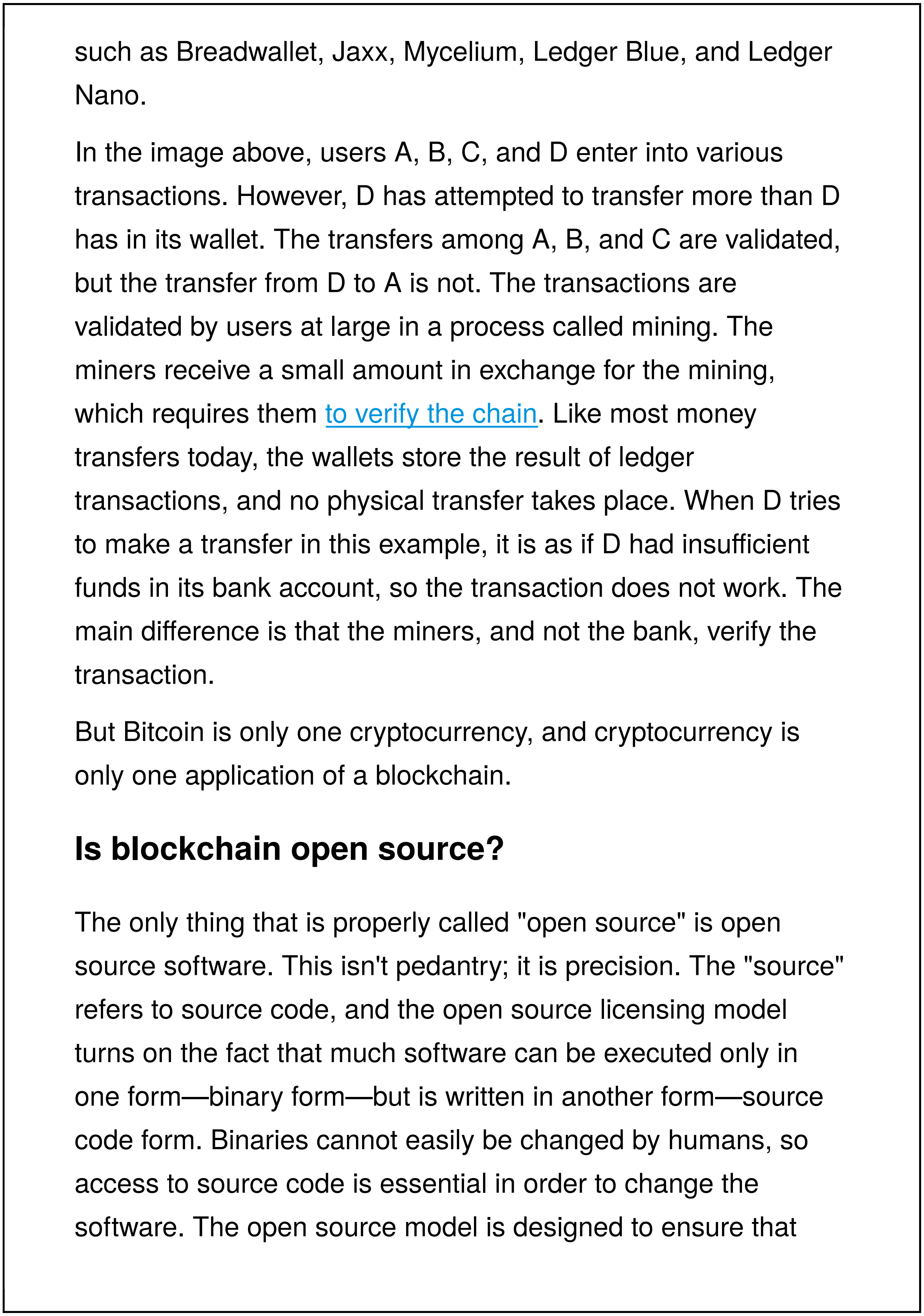
Bitcoins that could exist. As the demand for Bitcoin increased, the value increased, until eventually its price began to soar and has become quite volatile. Other cryptocurrencies do not necessarily follow the same rules and create scarcity by, for example, linking the currency to real-world items of value—similar to loyalty program points or scrip.

Bitcoin is implemented via blockchain technology.

If you own Bitcoins, they are stored in a wallet to which only you have access. Bitcoin wallets work somewhat like your email. After setting up an email address, you need software, usually called an email client, that enables you to send or receive emails. You can either download that software to a device or access an email client over the internet. To access your email, you employ a user name (an email ID) and a password. Similarly, Bitcoin wallets enable you to send or receive Bitcoins.

To access your wallet, you need two cryptographic keys: a public key and a private key. Public keys are known to everyone over the distributed network (like your email ID), but your private key is known only to you (like a password). When you access your Bitcoin wallet with your private key, you can transfer Bitcoins with anyone over the distributed network. No one can access your Bitcoin wallet without your private key. Consequently, every transaction you make will be recorded digitally in your Bitcoin wallet.

A Bitcoin wallet is one kind of digital wallet. There are other types of digital wallets, such as desktop wallets, cloud wallets, and mobile wallets. Various digital wallets work with Bitcoin,



such as Breadwallet, Jaxx, Mycelium, Ledger Blue, and Ledger Nano.

In the image above, users A, B, C, and D enter into various transactions. However, D has attempted to transfer more than D has in its wallet. The transfers among A, B, and C are validated, but the transfer from D to A is not. The transactions are validated by users at large in a process called mining. The miners receive a small amount in exchange for the mining, which requires them to verify the chain. Like most money transfers today, the wallets store the result of ledger transactions, and no physical transfer takes place. When D tries to make a transfer in this example, it is as if D had insufficient funds in its bank account, so the transaction does not work. The main difference is that the miners, and not the bank, verify the transaction.

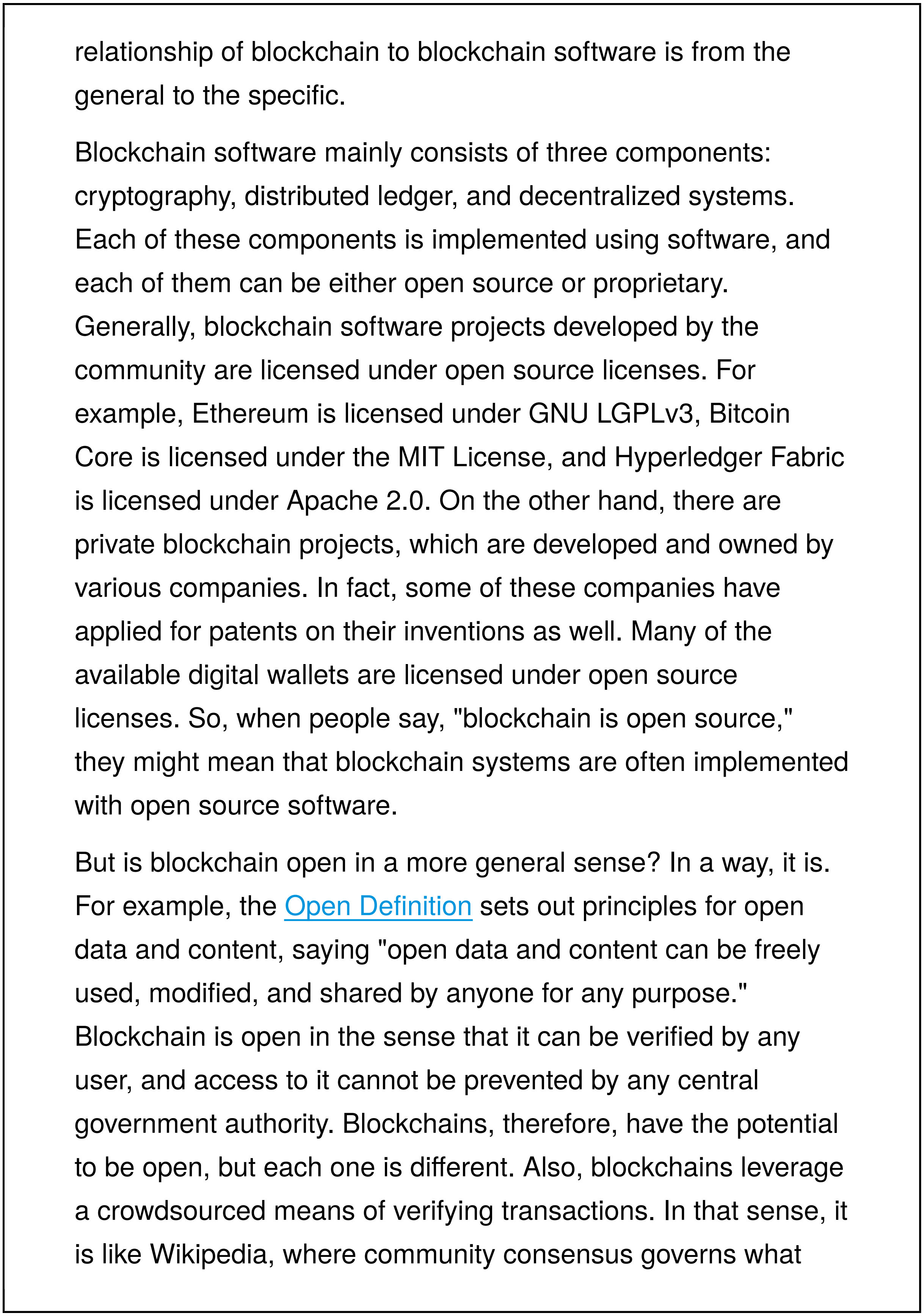
But Bitcoin is only one cryptocurrency, and cryptocurrency is only one application of a blockchain.

Is blockchain open source?

The only thing that is properly called "open source" is open source software. This isn't pedantry; it is precision. The "source" refers to source code, and the open source licensing model turns on the fact that much software can be executed only in one form—binary form—but is written in another form—source code form. Binaries cannot easily be changed by humans, so access to source code is essential in order to change the software. The open source model is designed to ensure that users of binary code have access to the source code for the binaries they are using so that they can examine, understand, fix, and improve that software. Most other things in the world do not have this quality—a dual nature where one form is readable and the other is not—so imposing the idea of open source on other things is awkward.

Although the term "open source" is often used to describe other things—from yoga to bioinformatics, to seders—the term can be misleading when applied more broadly than software. When people say something (other than software) is "open source," they usually mean one of two things: either it is available publicly or it is not subject to royalty-bearing patent claims. So, if you read or hear something is "open source," and that thing is not software, you need to ask more questions. These days, people often say "Bitcoin is open source" or "Blockchain is open source." So, what do they mean?

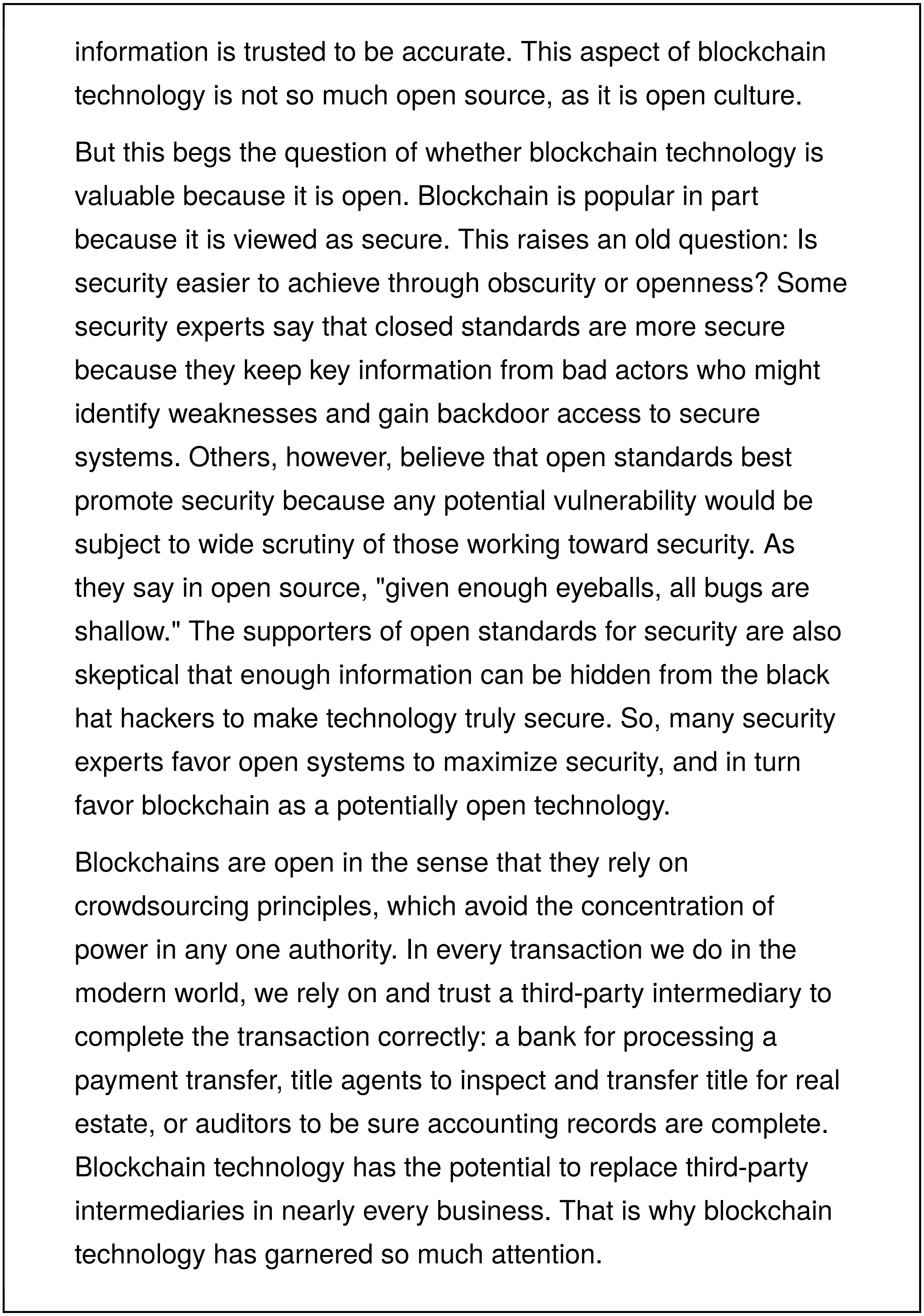
Blockchain is a technology or an ecosystem, and it is not the same as blockchain software. A blockchain is implemented via software, and there are various software projects that have been written to create and manage blockchains. In this sense, a blockchain is like a mathematical formula, like the quadratic equation or the formula to change Fahrenheit to Centigrade. You can write software to perform that function, but so could many others. If 100 coders each wrote a blockchain software program, they would all be different programs, perhaps written in different languages, with slight variations or large ones, but they would all have the same core functionality. Just as the relationship of blockchain to Bitcoin is from the general to the specific, the



relationship of blockchain to blockchain software is from the general to the specific.

Blockchain software mainly consists of three components: cryptography, distributed ledger, and decentralized systems. Each of these components is implemented using software, and each of them can be either open source or proprietary. Generally, blockchain software projects developed by the community are licensed under open source licenses. For example, Ethereum is licensed under GNU LGPLv3, Bitcoin Core is licensed under the MIT License, and Hyperledger Fabric is licensed under Apache 2.0. On the other hand, there are private blockchain projects, which are developed and owned by various companies. In fact, some of these companies have applied for patents on their inventions as well. Many of the available digital wallets are licensed under open source licenses. So, when people say, "blockchain is open source," they might mean that blockchain systems are often implemented with open source software.

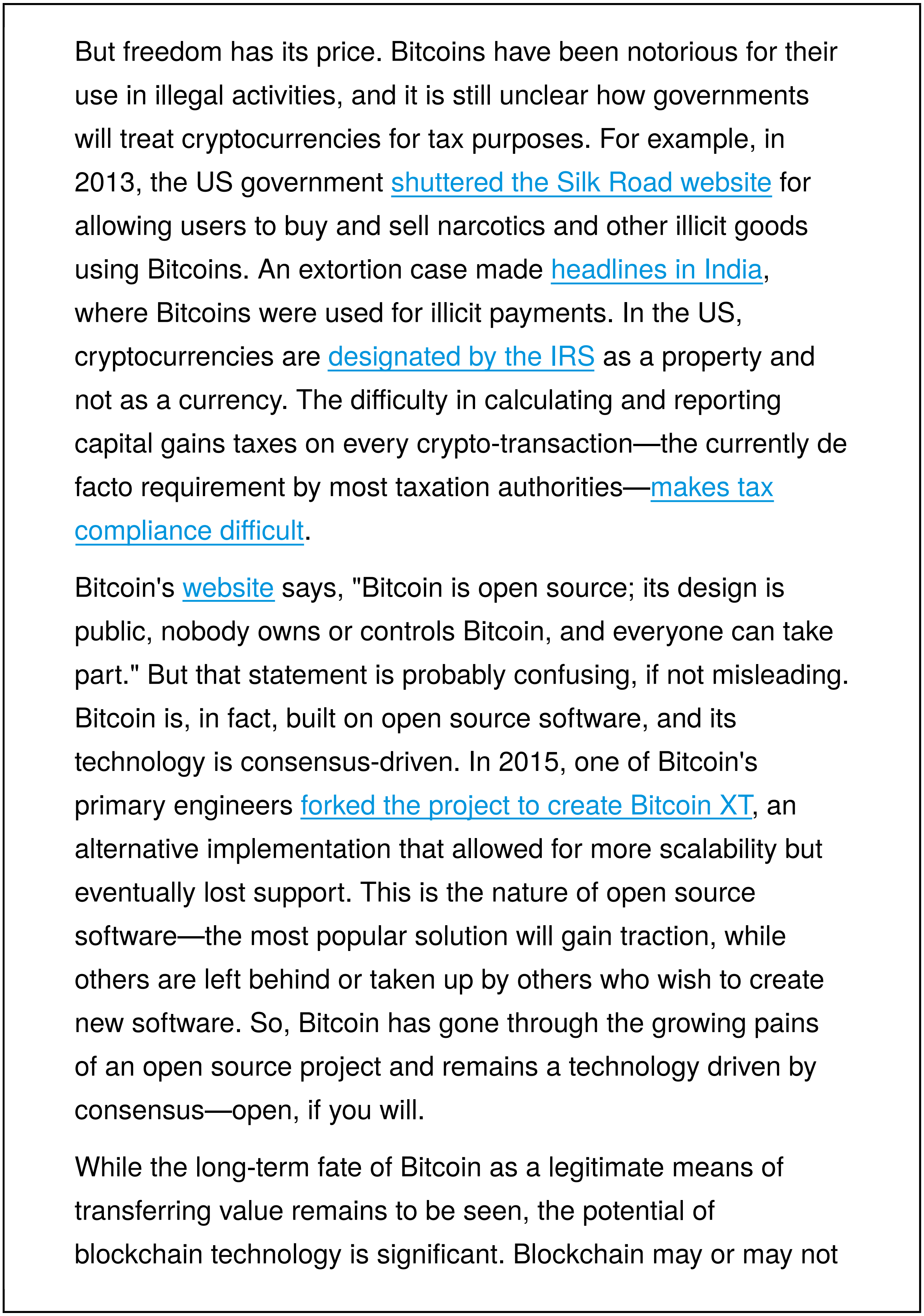
But is blockchain open in a more general sense? In a way, it is. For example, the Open Definition sets out principles for open data and content, saying "open data and content can be freely used, modified, and shared by anyone for any purpose." Blockchain is open in the sense that it can be verified by any user, and access to it cannot be prevented by any central government authority. Blockchains, therefore, have the potential to be open, but each one is different. Also, blockchains leverage a crowdsourced means of verifying transactions. In that sense, it is like Wikipedia, where community consensus governs what



information is trusted to be accurate. This aspect of blockchain technology is not so much open source, as it is open culture.

But this begs the question of whether blockchain technology is valuable because it is open. Blockchain is popular in part because it is viewed as secure. This raises an old question: Is security easier to achieve through obscurity or openness? Some security experts say that closed standards are more secure because they keep key information from bad actors who might identify weaknesses and gain backdoor access to secure systems. Others, however, believe that open standards best promote security because any potential vulnerability would be subject to wide scrutiny of those working toward security. As they say in open source, "given enough eyeballs, all bugs are shallow." The supporters of open standards for security are also skeptical that enough information can be hidden from the black hat hackers to make technology truly secure. So, many security experts favor open systems to maximize security, and in turn favor blockchain as a potentially open technology.

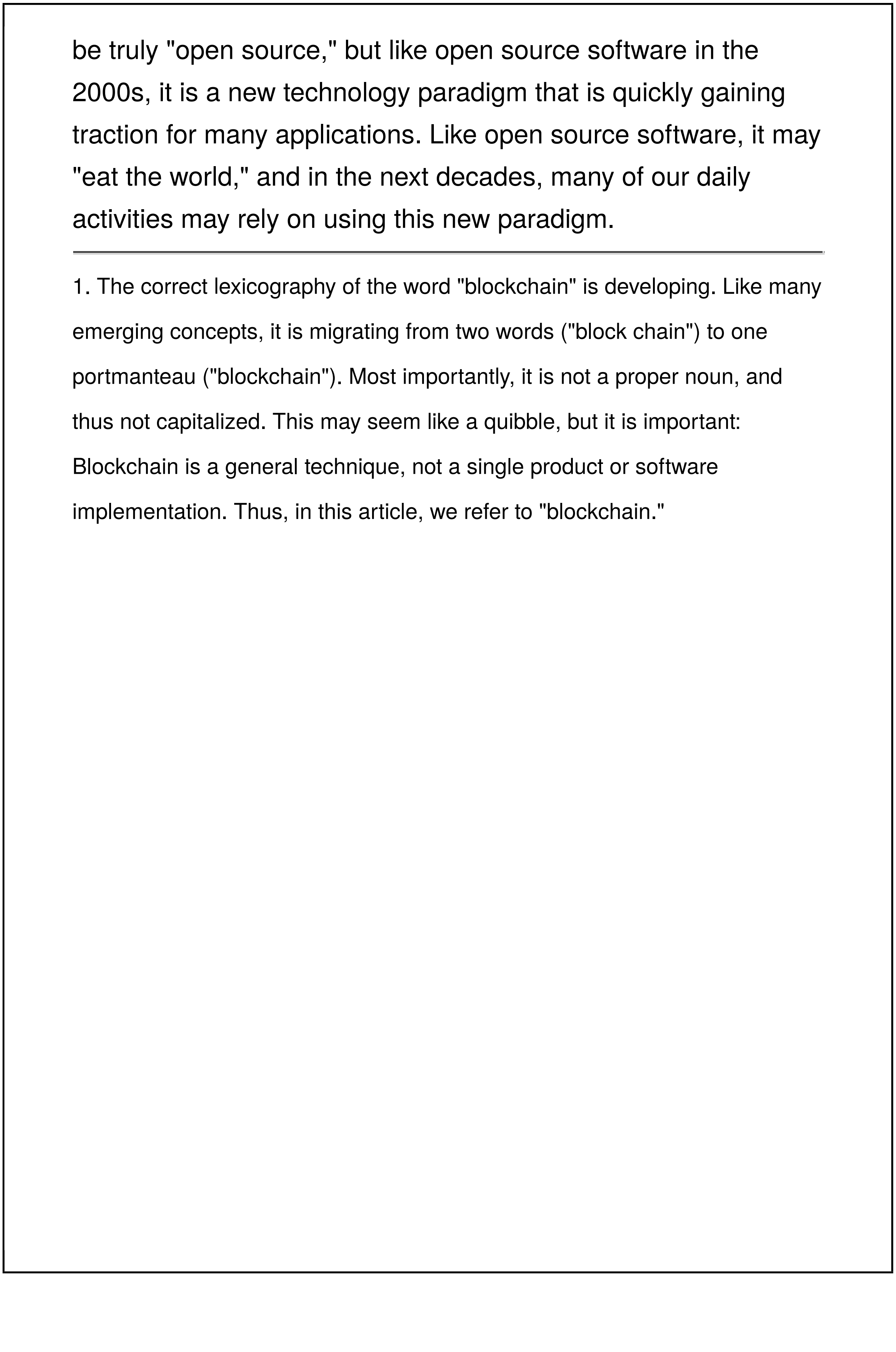
Blockchains are open in the sense that they rely on crowdsourcing principles, which avoid the concentration of power in any one authority. In every transaction we do in the modern world, we rely on and trust a third-party intermediary to complete the transaction correctly: a bank for processing a payment transfer, title agents to inspect and transfer title for real estate, or auditors to be sure accounting records are complete. Blockchain technology has the potential to replace third-party intermediaries in nearly every business. That is why blockchain technology has garnered so much attention.



But freedom has its price. Bitcoins have been notorious for their use in illegal activities, and it is still unclear how governments will treat cryptocurrencies for tax purposes. For example, in 2013, the US government shuttered the Silk Road website for allowing users to buy and sell narcotics and other illicit goods using Bitcoins. An extortion case made headlines in India, where Bitcoins were used for illicit payments. In the US, cryptocurrencies are designated by the IRS as a property and not as a currency. The difficulty in calculating and reporting capital gains taxes on every crypto-transaction—the currently de facto requirement by most taxation authorities—makes tax compliance difficult.

Bitcoin's website says, "Bitcoin is open source; its design is public, nobody owns or controls Bitcoin, and everyone can take part." But that statement is probably confusing, if not misleading. Bitcoin is, in fact, built on open source software, and its technology is consensus-driven. In 2015, one of Bitcoin's primary engineers forked the project to create Bitcoin XT, an alternative implementation that allowed for more scalability but eventually lost support. This is the nature of open source software—the most popular solution will gain traction, while others are left behind or taken up by others who wish to create new software. So, Bitcoin has gone through the growing pains of an open source project and remains a technology driven by consensus—open, if you will.

While the long-term fate of Bitcoin as a legitimate means of transferring value remains to be seen, the potential of blockchain technology is significant. Blockchain may or may not

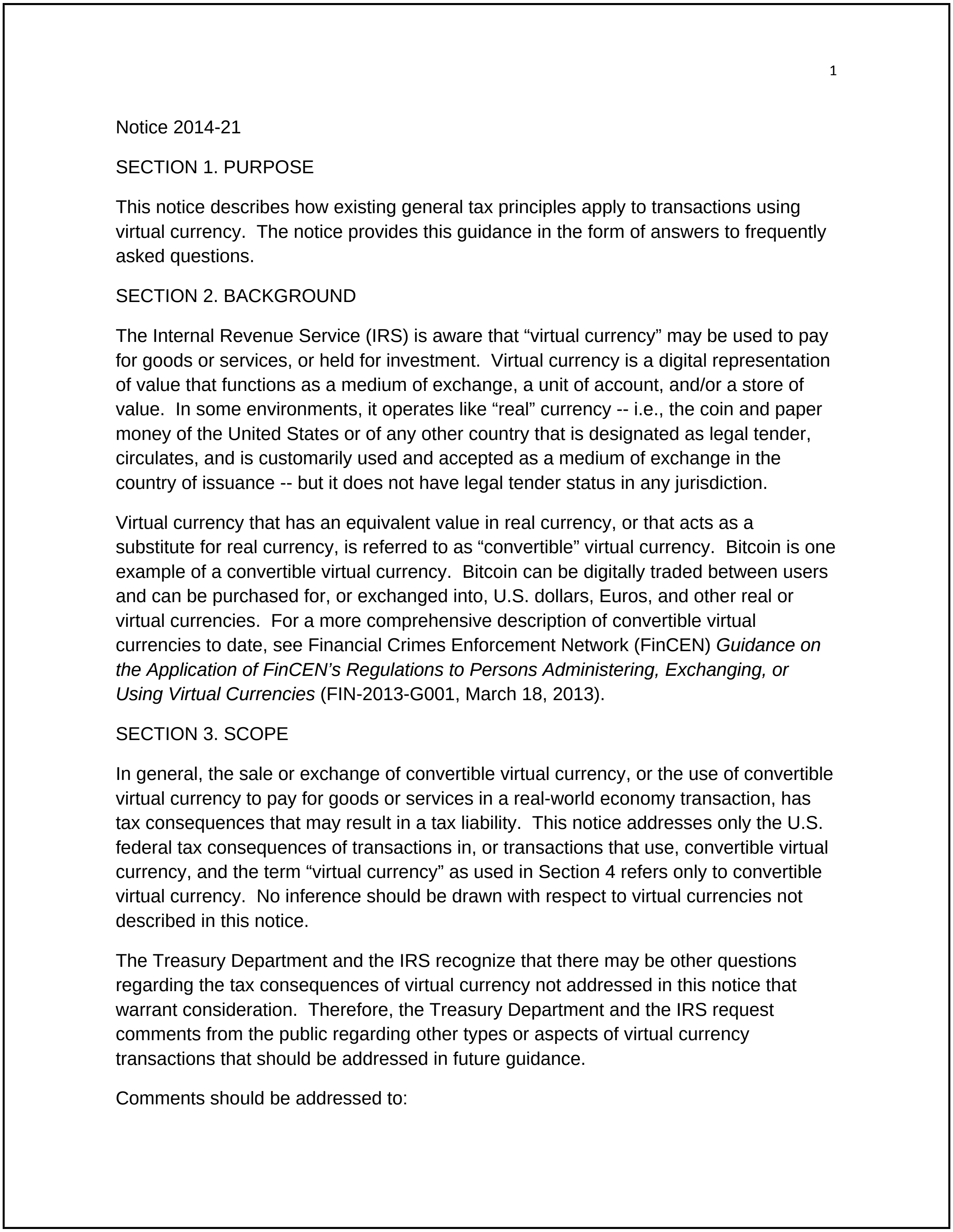


be truly "open source," but like open source software in the 2000s, it is a new technology paradigm that is quickly gaining traction for many applications. Like open source software, it may "eat the world," and in the next decades, many of our daily activities may rely on using this new paradigm.

[1] The correct lexicography of the word "blockchain" is developing. Like many emerging concepts, it is migrating from two words ("block chain") to one portmanteau ("blockchain"). Most importantly, it is not a proper noun, and thus not capitalized. This may seem like a quibble, but it is important: Blockchain is a general technique, not a single product or software implementation. Thus, in this article, we refer to "blockchain."

Chapter 19

IRS Notice 2014-21



Notice 2014-21

SECTION 1. PURPOSE

This notice describes how existing general tax principles apply to transactions using virtual currency. The notice provides this guidance in the form of answers to frequently asked questions.

SECTION 2. BACKGROUND

The Internal Revenue Service (IRS) is aware that “virtual currency” may be used to pay for goods or services, or held for investment. Virtual currency is a digital representation of value that functions as a medium of exchange, a unit of account, and/or a store of value. In some environments, it operates like “real” currency -- i.e., the coin and paper money of the United States or of any other country that is designated as legal tender, circulates, and is customarily used and accepted as a medium of exchange in the country of issuance -- but it does not have legal tender status in any jurisdiction.

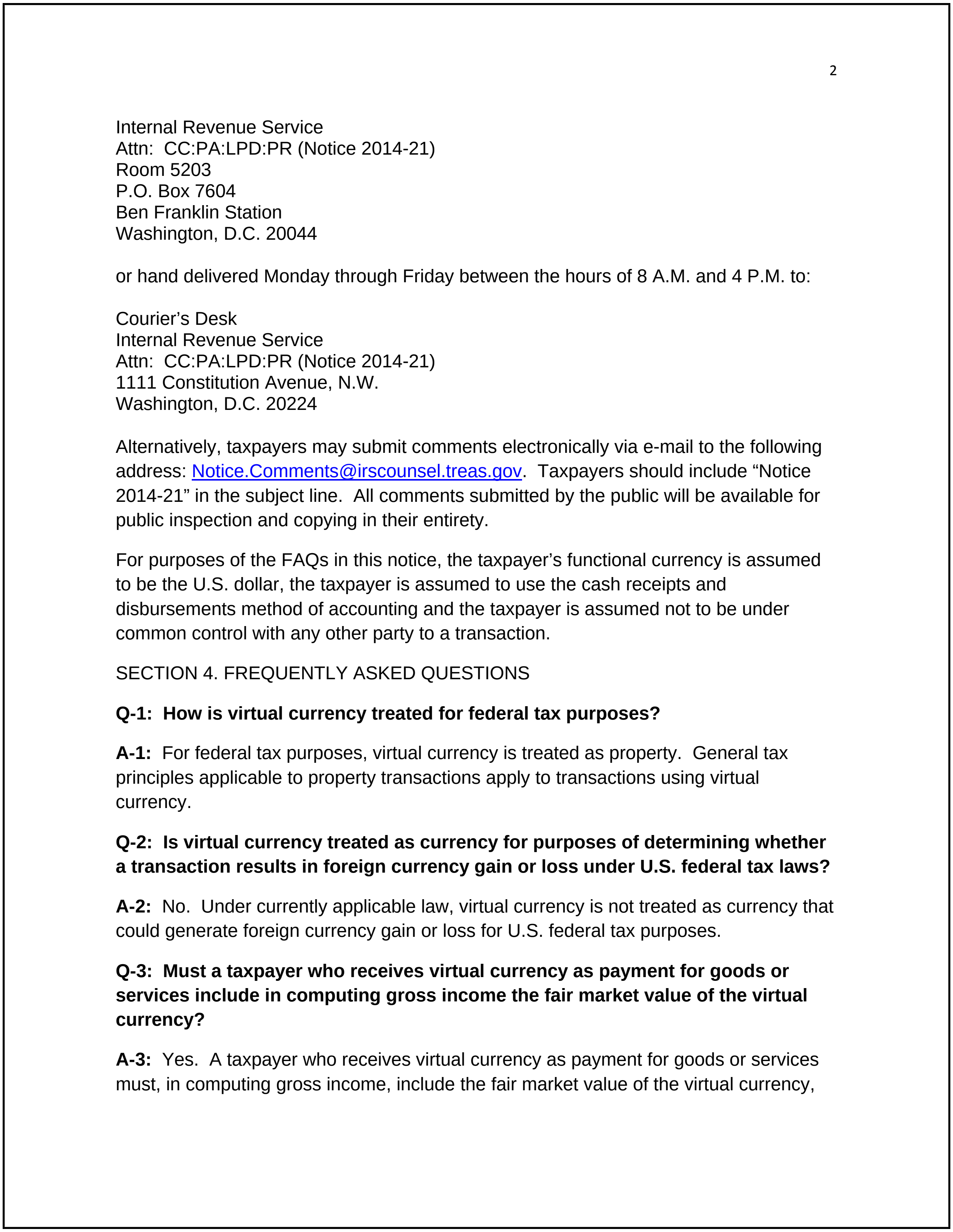
Virtual currency that has an equivalent value in real currency, or that acts as a substitute for real currency, is referred to as “convertible” virtual currency. Bitcoin is one example of a convertible virtual currency. Bitcoin can be digitally traded between users and can be purchased for, or exchanged into, U.S. dollars, Euros, and other real or virtual currencies. For a more comprehensive description of convertible virtual currencies to date, see Financial Crimes Enforcement Network (FinCEN) *Guidance on the Application of FinCEN’s Regulations to Persons Administering, Exchanging, or Using Virtual Currencies* (FIN-2013-G001, March 18, 2013).

SECTION 3. SCOPE

In general, the sale or exchange of convertible virtual currency, or the use of convertible virtual currency to pay for goods or services in a real-world economy transaction, has tax consequences that may result in a tax liability. This notice addresses only the U.S. federal tax consequences of transactions in, or transactions that use, convertible virtual currency, and the term “virtual currency” as used in Section 4 refers only to convertible virtual currency. No inference should be drawn with respect to virtual currencies not described in this notice.

The Treasury Department and the IRS recognize that there may be other questions regarding the tax consequences of virtual currency not addressed in this notice that warrant consideration. Therefore, the Treasury Department and the IRS request comments from the public regarding other types or aspects of virtual currency transactions that should be addressed in future guidance.

Comments should be addressed to:



Internal Revenue Service
Attn: CC:PA:LPD:PR (Notice 2014-21)
Room 5203
P.O. Box 7604
Ben Franklin Station
Washington, D.C. 20044

or hand delivered Monday through Friday between the hours of 8 A.M. and 4 P.M. to:

Courier's Desk
Internal Revenue Service
Attn: CC:PA:LPD:PR (Notice 2014-21)
1111 Constitution Avenue, N.W.
Washington, D.C. 20224

Alternatively, taxpayers may submit comments electronically via e-mail to the following address: Notice.Comments@irscounsel.treas.gov. Taxpayers should include "Notice 2014-21" in the subject line. All comments submitted by the public will be available for public inspection and copying in their entirety.

For purposes of the FAQs in this notice, the taxpayer's functional currency is assumed to be the U.S. dollar, the taxpayer is assumed to use the cash receipts and disbursements method of accounting and the taxpayer is assumed not to be under common control with any other party to a transaction.

SECTION 4. FREQUENTLY ASKED QUESTIONS

Q-1: How is virtual currency treated for federal tax purposes?

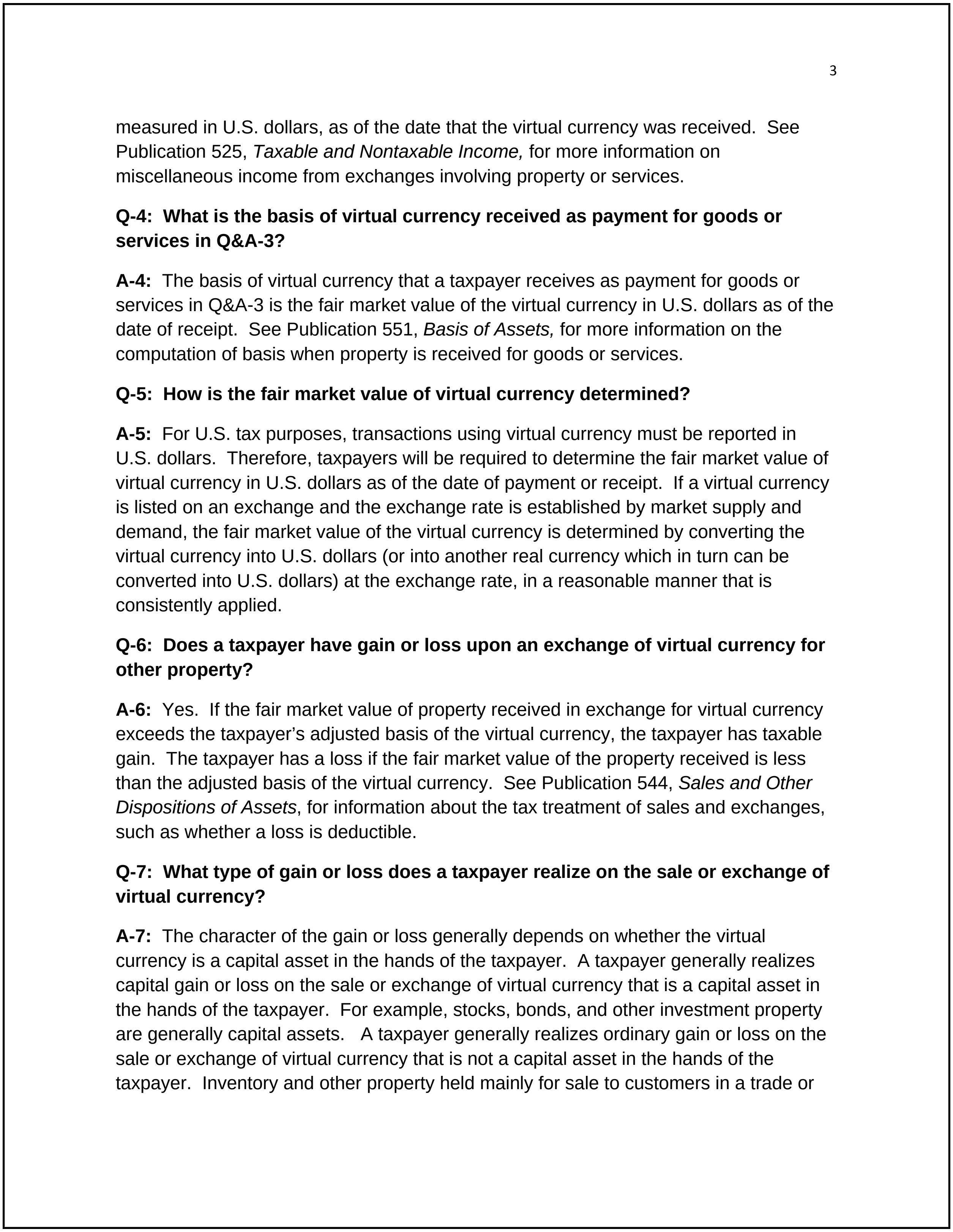
A-1: For federal tax purposes, virtual currency is treated as property. General tax principles applicable to property transactions apply to transactions using virtual currency.

Q-2: Is virtual currency treated as currency for purposes of determining whether a transaction results in foreign currency gain or loss under U.S. federal tax laws?

A-2: No. Under currently applicable law, virtual currency is not treated as currency that could generate foreign currency gain or loss for U.S. federal tax purposes.

Q-3: Must a taxpayer who receives virtual currency as payment for goods or services include in computing gross income the fair market value of the virtual currency?

A-3: Yes. A taxpayer who receives virtual currency as payment for goods or services must, in computing gross income, include the fair market value of the virtual currency,



measured in U.S. dollars, as of the date that the virtual currency was received. See Publication 525, *Taxable and Nontaxable Income,* for more information on miscellaneous income from exchanges involving property or services.

Q-4: What is the basis of virtual currency received as payment for goods or services in Q&A-3?

A-4: The basis of virtual currency that a taxpayer receives as payment for goods or services in Q&A-3 is the fair market value of the virtual currency in U.S. dollars as of the date of receipt. See Publication 551, *Basis of Assets,* for more information on the computation of basis when property is received for goods or services.

Q-5: How is the fair market value of virtual currency determined?

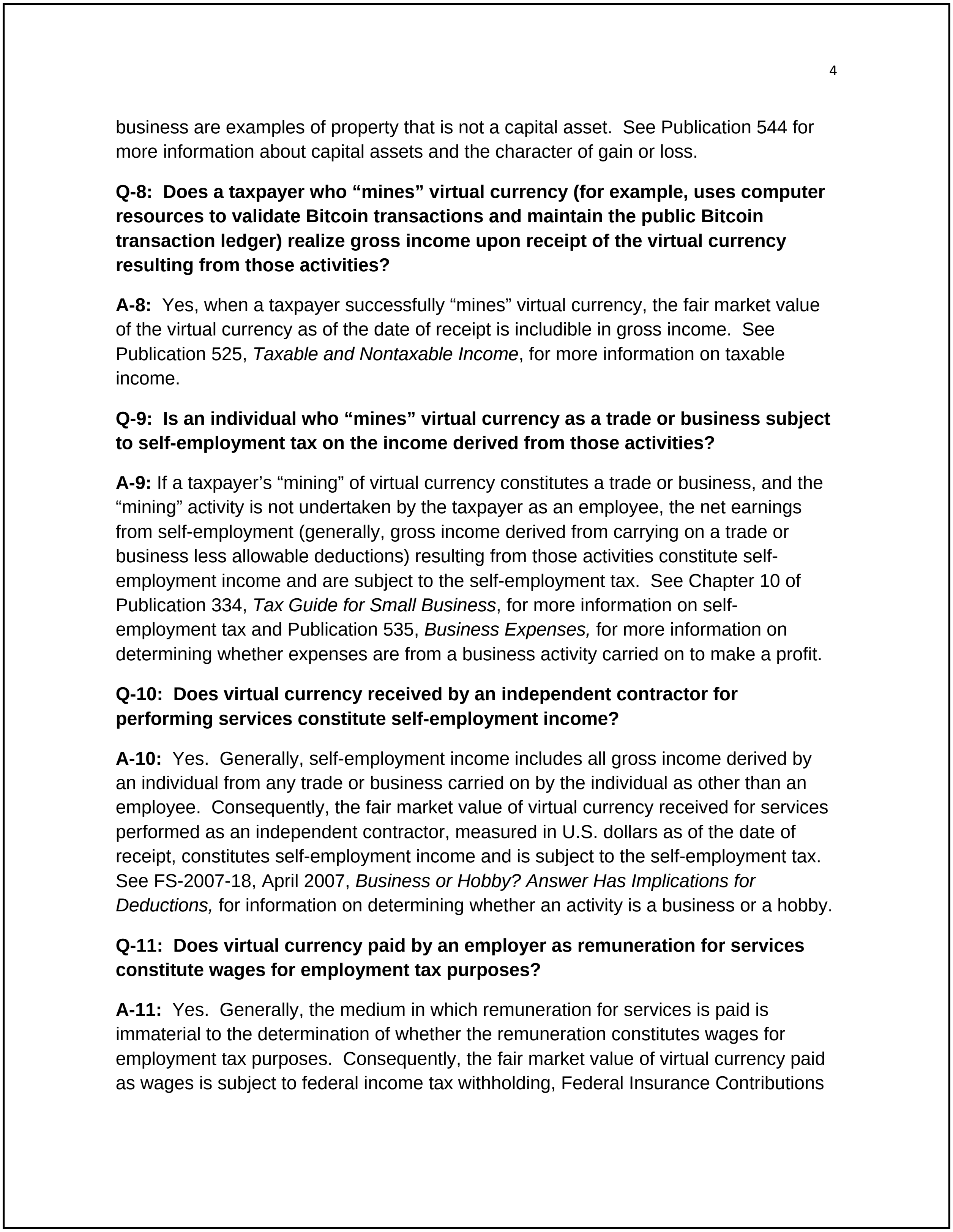
A-5: For U.S. tax purposes, transactions using virtual currency must be reported in U.S. dollars. Therefore, taxpayers will be required to determine the fair market value of virtual currency in U.S. dollars as of the date of payment or receipt. If a virtual currency is listed on an exchange and the exchange rate is established by market supply and demand, the fair market value of the virtual currency is determined by converting the virtual currency into U.S. dollars (or into another real currency which in turn can be converted into U.S. dollars) at the exchange rate, in a reasonable manner that is consistently applied.

Q-6: Does a taxpayer have gain or loss upon an exchange of virtual currency for other property?

A-6: Yes. If the fair market value of property received in exchange for virtual currency exceeds the taxpayer's adjusted basis of the virtual currency, the taxpayer has taxable gain. The taxpayer has a loss if the fair market value of the property received is less than the adjusted basis of the virtual currency. See Publication 544, *Sales and Other Dispositions of Assets*, for information about the tax treatment of sales and exchanges, such as whether a loss is deductible.

Q-7: What type of gain or loss does a taxpayer realize on the sale or exchange of virtual currency?

A-7: The character of the gain or loss generally depends on whether the virtual currency is a capital asset in the hands of the taxpayer. A taxpayer generally realizes capital gain or loss on the sale or exchange of virtual currency that is a capital asset in the hands of the taxpayer. For example, stocks, bonds, and other investment property are generally capital assets. A taxpayer generally realizes ordinary gain or loss on the sale or exchange of virtual currency that is not a capital asset in the hands of the taxpayer. Inventory and other property held mainly for sale to customers in a trade or



business are examples of property that is not a capital asset. See Publication 544 for more information about capital assets and the character of gain or loss.

Q-8: Does a taxpayer who "mines" virtual currency (for example, uses computer resources to validate Bitcoin transactions and maintain the public Bitcoin transaction ledger) realize gross income upon receipt of the virtual currency resulting from those activities?

A-8: Yes, when a taxpayer successfully "mines" virtual currency, the fair market value of the virtual currency as of the date of receipt is includible in gross income. See Publication 525, *Taxable and Nontaxable Income*, for more information on taxable income.

Q-9: Is an individual who "mines" virtual currency as a trade or business subject to self-employment tax on the income derived from those activities?

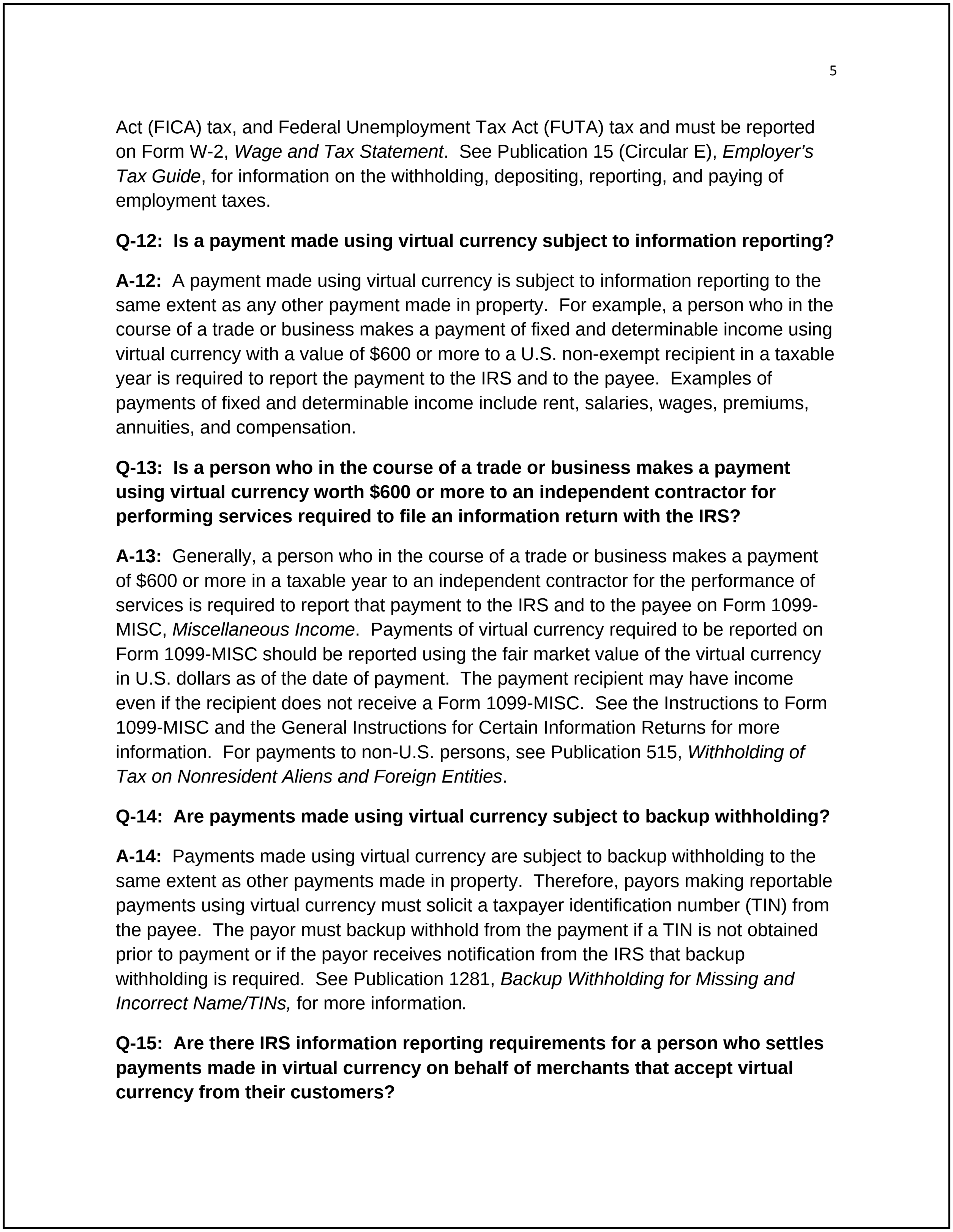
A-9: If a taxpayer's "mining" of virtual currency constitutes a trade or business, and the "mining" activity is not undertaken by the taxpayer as an employee, the net earnings from self-employment (generally, gross income derived from carrying on a trade or business less allowable deductions) resulting from those activities constitute self-employment income and are subject to the self-employment tax. See Chapter 10 of Publication 334, *Tax Guide for Small Business*, for more information on self-employment tax and Publication 535, *Business Expenses,* for more information on determining whether expenses are from a business activity carried on to make a profit.

Q-10: Does virtual currency received by an independent contractor for performing services constitute self-employment income?

A-10: Yes. Generally, self-employment income includes all gross income derived by an individual from any trade or business carried on by the individual as other than an employee. Consequently, the fair market value of virtual currency received for services performed as an independent contractor, measured in U.S. dollars as of the date of receipt, constitutes self-employment income and is subject to the self-employment tax. See FS-2007-18, April 2007, *Business or Hobby? Answer Has Implications for Deductions,* for information on determining whether an activity is a business or a hobby.

Q-11: Does virtual currency paid by an employer as remuneration for services constitute wages for employment tax purposes?

A-11: Yes. Generally, the medium in which remuneration for services is paid is immaterial to the determination of whether the remuneration constitutes wages for employment tax purposes. Consequently, the fair market value of virtual currency paid as wages is subject to federal income tax withholding, Federal Insurance Contributions



Act (FICA) tax, and Federal Unemployment Tax Act (FUTA) tax and must be reported on Form W-2, *Wage and Tax Statement*. See Publication 15 (Circular E), *Employer's Tax Guide*, for information on the withholding, depositing, reporting, and paying of employment taxes.

Q-12: Is a payment made using virtual currency subject to information reporting?

A-12: A payment made using virtual currency is subject to information reporting to the same extent as any other payment made in property. For example, a person who in the course of a trade or business makes a payment of fixed and determinable income using virtual currency with a value of $600 or more to a U.S. non-exempt recipient in a taxable year is required to report the payment to the IRS and to the payee. Examples of payments of fixed and determinable income include rent, salaries, wages, premiums, annuities, and compensation.

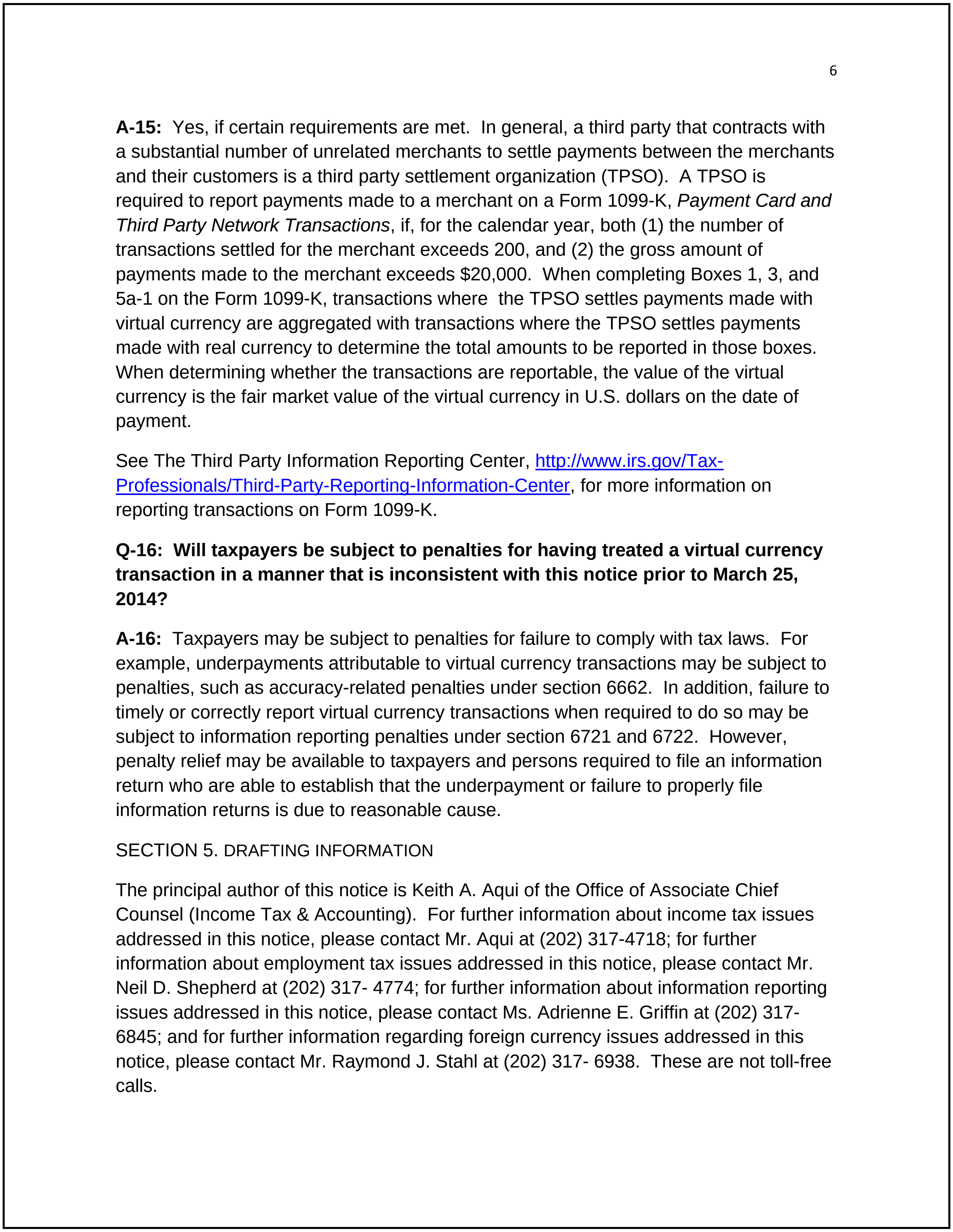
Q-13: Is a person who in the course of a trade or business makes a payment using virtual currency worth $600 or more to an independent contractor for performing services required to file an information return with the IRS?

A-13: Generally, a person who in the course of a trade or business makes a payment of $600 or more in a taxable year to an independent contractor for the performance of services is required to report that payment to the IRS and to the payee on Form 1099-MISC, *Miscellaneous Income*. Payments of virtual currency required to be reported on Form 1099-MISC should be reported using the fair market value of the virtual currency in U.S. dollars as of the date of payment. The payment recipient may have income even if the recipient does not receive a Form 1099-MISC. See the Instructions to Form 1099-MISC and the General Instructions for Certain Information Returns for more information. For payments to non-U.S. persons, see Publication 515, *Withholding of Tax on Nonresident Aliens and Foreign Entities*.

Q-14: Are payments made using virtual currency subject to backup withholding?

A-14: Payments made using virtual currency are subject to backup withholding to the same extent as other payments made in property. Therefore, payors making reportable payments using virtual currency must solicit a taxpayer identification number (TIN) from the payee. The payor must backup withhold from the payment if a TIN is not obtained prior to payment or if the payor receives notification from the IRS that backup withholding is required. See Publication 1281, *Backup Withholding for Missing and Incorrect Name/TINs,* for more information*.*

Q-15: Are there IRS information reporting requirements for a person who settles payments made in virtual currency on behalf of merchants that accept virtual currency from their customers?



A-15: Yes, if certain requirements are met. In general, a third party that contracts with a substantial number of unrelated merchants to settle payments between the merchants and their customers is a third party settlement organization (TPSO). A TPSO is required to report payments made to a merchant on a Form 1099-K, *Payment Card and Third Party Network Transactions*, if, for the calendar year, both (1) the number of transactions settled for the merchant exceeds 200, and (2) the gross amount of payments made to the merchant exceeds $20,000. When completing Boxes 1, 3, and 5a-1 on the Form 1099-K, transactions where the TPSO settles payments made with virtual currency are aggregated with transactions where the TPSO settles payments made with real currency to determine the total amounts to be reported in those boxes. When determining whether the transactions are reportable, the value of the virtual currency is the fair market value of the virtual currency in U.S. dollars on the date of payment.

See The Third Party Information Reporting Center, http://www.irs.gov/Tax-Professionals/Third-Party-Reporting-Information-Center, for more information on reporting transactions on Form 1099-K.

Q-16: Will taxpayers be subject to penalties for having treated a virtual currency transaction in a manner that is inconsistent with this notice prior to March 25, 2014?

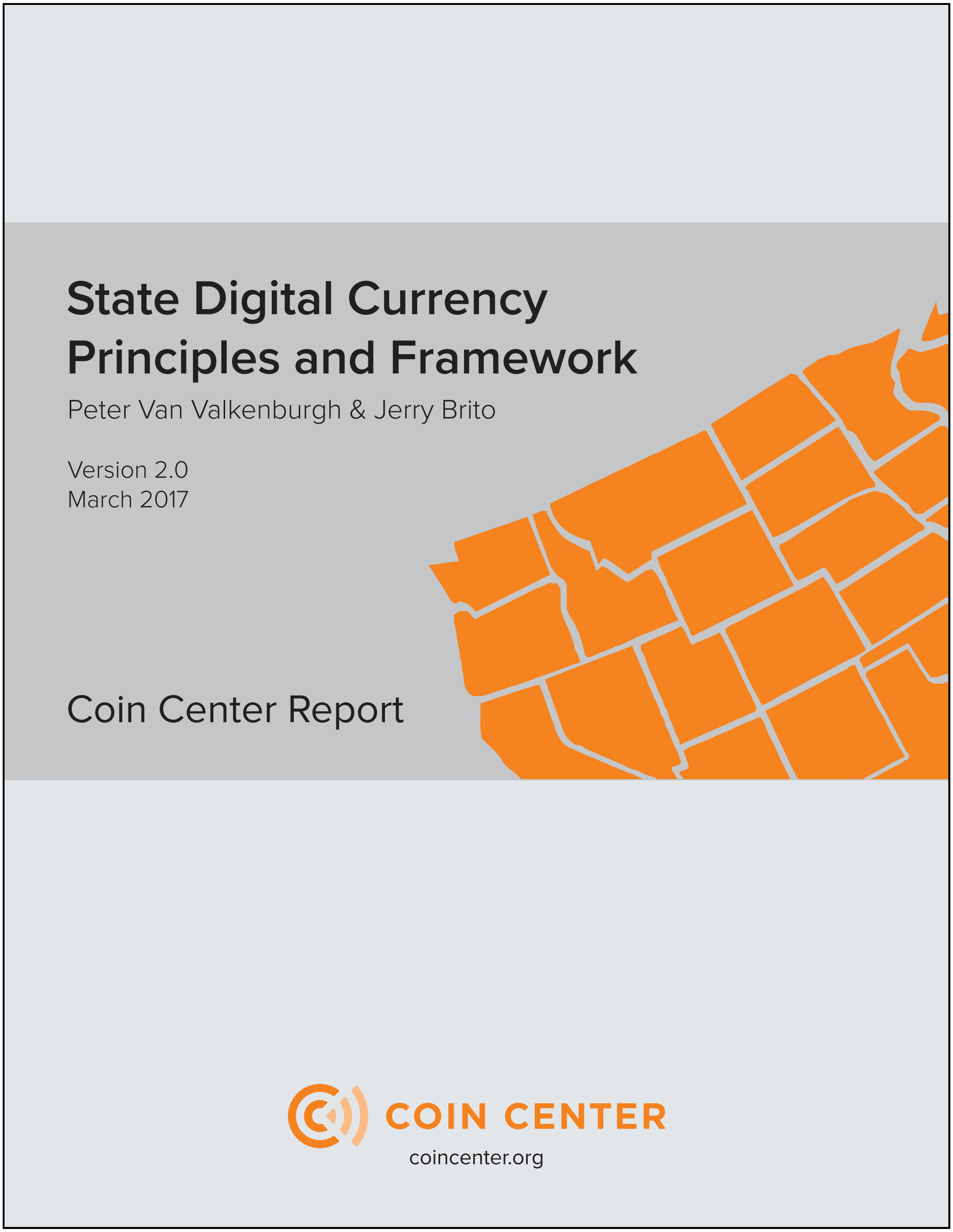
A-16: Taxpayers may be subject to penalties for failure to comply with tax laws. For example, underpayments attributable to virtual currency transactions may be subject to penalties, such as accuracy-related penalties under section 6662. In addition, failure to timely or correctly report virtual currency transactions when required to do so may be subject to information reporting penalties under section 6721 and 6722. However, penalty relief may be available to taxpayers and persons required to file an information return who are able to establish that the underpayment or failure to properly file information returns is due to reasonable cause.

SECTION 5. DRAFTING INFORMATION

The principal author of this notice is Keith A. Aqui of the Office of Associate Chief Counsel (Income Tax & Accounting). For further information about income tax issues addressed in this notice, please contact Mr. Aqui at (202) 317-4718; for further information about employment tax issues addressed in this notice, please contact Mr. Neil D. Shepherd at (202) 317- 4774; for further information about information reporting issues addressed in this notice, please contact Ms. Adrienne E. Griffin at (202) 317-6845; and for further information regarding foreign currency issues addressed in this notice, please contact Mr. Raymond J. Stahl at (202) 317- 6938. These are not toll-free calls.

Chapter 20

State Digital Currency Principles and Framework (Van Valkenburgh and Brito)



State Digital Currency Principles and Framework

Peter Van Valkenburgh & Jerry Brito

Version 2.0
March 2017

Coin Center Report

COIN CENTER
coincenter.org

Peter Van Valkenburgh and Jerry Brito, *State Virtual Currency Principles and Framework v2.0*, Coin Center Report, Mar. 2017, available at https://coincenter.org/entry/state-digital-currency-principles-and-framework

Abstract

States have begun to look at how virtual currencies, such as Bitcoin, and the businesses that utilize them to provide consumer products, interact with money transmission and consumer protection policy. This report reviews the approaches taken by the several States thus far, and offers model language for a *sui generis* statute or amendment to a money transmission statute. It is not a draft or model bill in full. Instead, language is offered for the essential components of any virtual currency law: Who must be licensed? How do you define “control” of customer virtual currency? How are startups encouraged while still protecting consumers? How is solvency guaranteed?

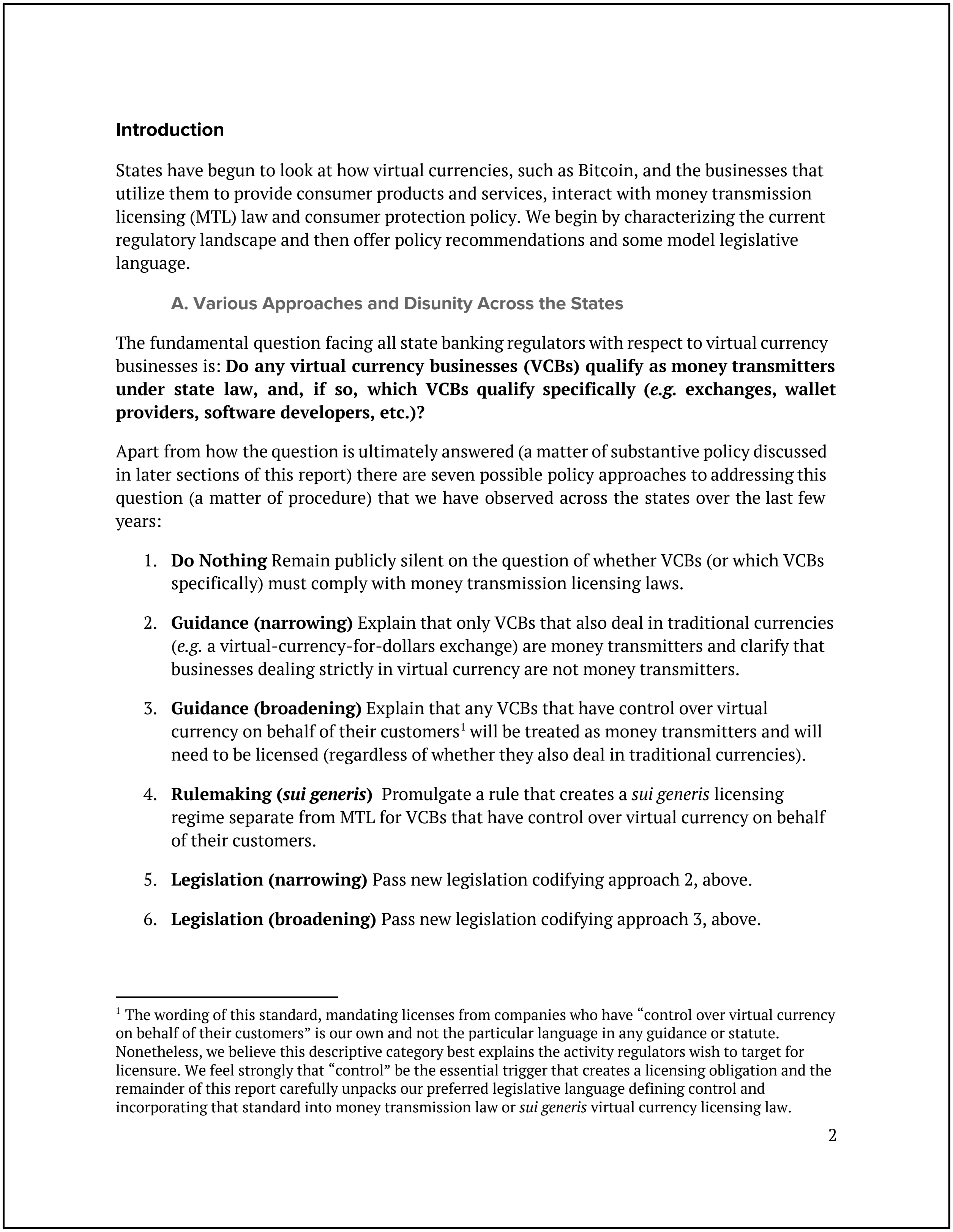
Author

Peter Van Valkenburgh
Director of Research
Coin Center
peter@coincenter.org

Jerry Brito
Executive Director
Coin Center
jerry@coincenter.org

About Coin Center

Coin Center is a non-profit research and advocacy center focused on the public policy issues facing cryptocurrency technologies such as Bitcoin. Our mission is to build a better understanding of these technologies and to promote a regulatory climate that preserves the freedom to innovate using blockchain technologies. We do this by producing and publishing policy research from respected academics and experts, educating policymakers and the media about blockchain technology, and by engaging in advocacy for sound public policy.



Introduction

States have begun to look at how virtual currencies, such as Bitcoin, and the businesses that utilize them to provide consumer products and services, interact with money transmission licensing (MTL) law and consumer protection policy. We begin by characterizing the current regulatory landscape and then offer policy recommendations and some model legislative language.

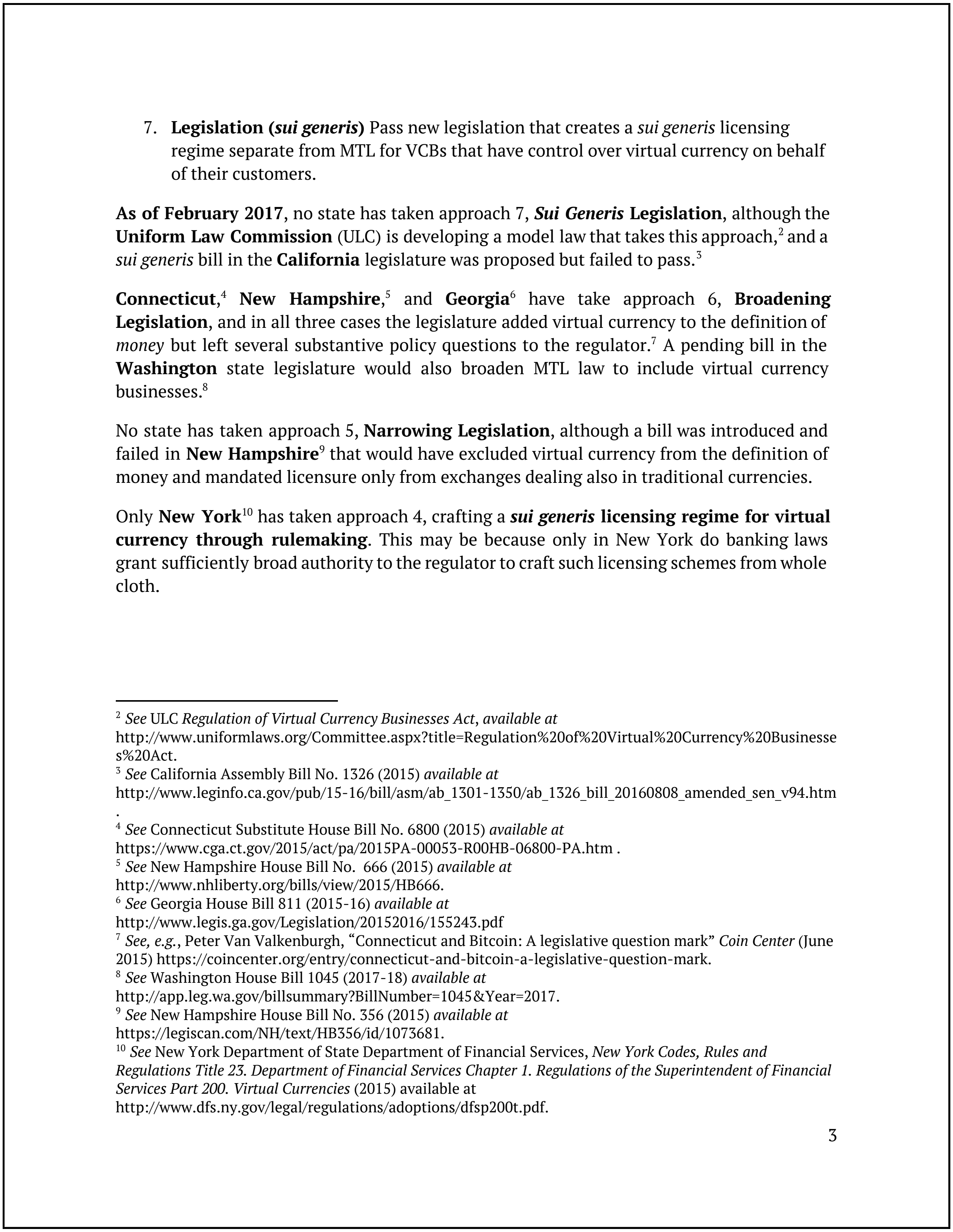
A. Various Approaches and Disunity Across the States

The fundamental question facing all state banking regulators with respect to virtual currency businesses is: **Do any virtual currency businesses (VCBs) qualify as money transmitters under state law, and, if so, which VCBs qualify specifically (*e.g.* exchanges, wallet providers, software developers, etc.)?**

Apart from how the question is ultimately answered (a matter of substantive policy discussed in later sections of this report) there are seven possible policy approaches to addressing this question (a matter of procedure) that we have observed across the states over the last few years:

1. **Do Nothing** Remain publicly silent on the question of whether VCBs (or which VCBs specifically) must comply with money transmission licensing laws.
2. **Guidance (narrowing)** Explain that only VCBs that also deal in traditional currencies (*e.g.* a virtual-currency-for-dollars exchange) are money transmitters and clarify that businesses dealing strictly in virtual currency are not money transmitters.
3. **Guidance (broadening)** Explain that any VCBs that have control over virtual currency on behalf of their customers[1] will be treated as money transmitters and will need to be licensed (regardless of whether they also deal in traditional currencies).
4. **Rulemaking (*sui generis*)** Promulgate a rule that creates a *sui generis* licensing regime separate from MTL for VCBs that have control over virtual currency on behalf of their customers.
5. **Legislation (narrowing)** Pass new legislation codifying approach 2, above.
6. **Legislation (broadening)** Pass new legislation codifying approach 3, above.

[1] The wording of this standard, mandating licenses from companies who have "control over virtual currency on behalf of their customers" is our own and not the particular language in any guidance or statute. Nonetheless, we believe this descriptive category best explains the activity regulators wish to target for licensure. We feel strongly that "control" be the essential trigger that creates a licensing obligation and the remainder of this report carefully unpacks our preferred legislative language defining control and incorporating that standard into money transmission law or *sui generis* virtual currency licensing law.



7. **Legislation (*sui generis*)** Pass new legislation that creates a *sui generis* licensing regime separate from MTL for VCBs that have control over virtual currency on behalf of their customers.

As of February 2017, no state has taken approach 7, ***Sui Generis* Legislation**, although the **Uniform Law Commission** (ULC) is developing a model law that takes this approach,[2] and a *sui generis* bill in the **California** legislature was proposed but failed to pass.[3]

Connecticut,[4] **New Hampshire**,[5] and **Georgia**[6] have take approach 6, **Broadening Legislation**, and in all three cases the legislature added virtual currency to the definition of *money* but left several substantive policy questions to the regulator.[7] A pending bill in the **Washington** state legislature would also broaden MTL law to include virtual currency businesses.[8]

No state has taken approach 5, **Narrowing Legislation**, although a bill was introduced and failed in **New Hampshire**[9] that would have excluded virtual currency from the definition of money and mandated licensure only from exchanges dealing also in traditional currencies.

Only **New York**[10] has taken approach 4, crafting a ***sui generis* licensing regime for virtual currency through rulemaking**. This may be because only in New York do banking laws grant sufficiently broad authority to the regulator to craft such licensing schemes from whole cloth.

[2] *See* ULC *Regulation of Virtual Currency Businesses Act*, *available at* http://www.uniformlaws.org/Committee.aspx?title=Regulation%20of%20Virtual%20Currency%20Businesses%20Act.

[3] *See* California Assembly Bill No. 1326 (2015) *available at* http://www.leginfo.ca.gov/pub/15-16/bill/asm/ab_1301-1350/ab_1326_bill_20160808_amended_sen_v94.htm .

[4] *See* Connecticut Substitute House Bill No. 6800 (2015) *available at* https://www.cga.ct.gov/2015/act/pa/2015PA-00053-R00HB-06800-PA.htm .

[5] *See* New Hampshire House Bill No. 666 (2015) *available at* http://www.nhliberty.org/bills/view/2015/HB666.

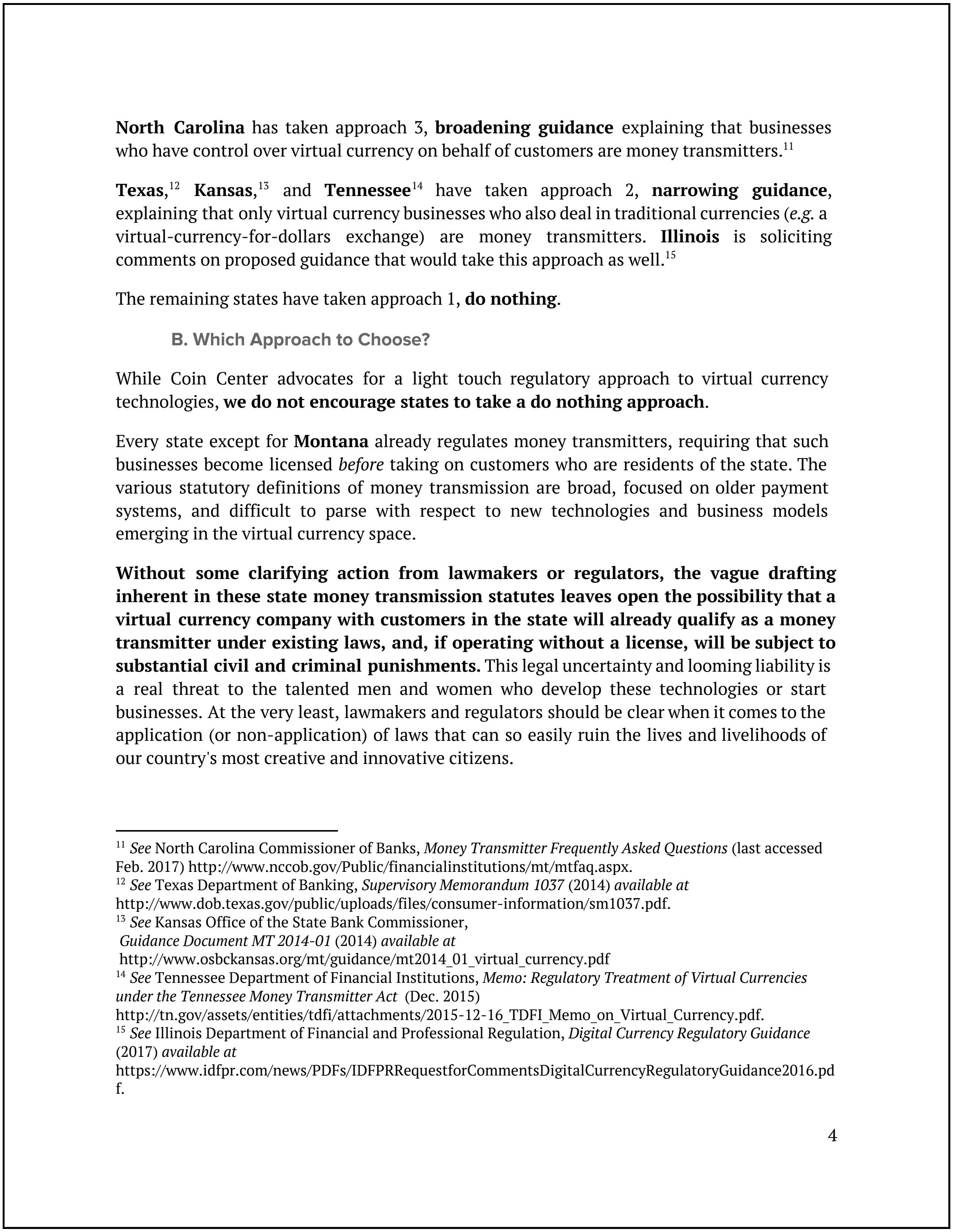
[6] *See* Georgia House Bill 811 (2015-16) *available at* http://www.legis.ga.gov/Legislation/20152016/155243.pdf

[7] *See, e.g.*, Peter Van Valkenburgh, "Connecticut and Bitcoin: A legislative question mark" *Coin Center* (June 2015) https://coincenter.org/entry/connecticut-and-bitcoin-a-legislative-question-mark.

[8] *See* Washington House Bill 1045 (2017-18) *available at* http://app.leg.wa.gov/billsummary?BillNumber=1045&Year=2017.

[9] *See* New Hampshire House Bill No. 356 (2015) *available at* https://legiscan.com/NH/text/HB356/id/1073681.

[10] *See* New York Department of State Department of Financial Services, *New York Codes, Rules and Regulations Title 23. Department of Financial Services Chapter 1. Regulations of the Superintendent of Financial Services Part 200. Virtual Currencies* (2015) available at http://www.dfs.ny.gov/legal/regulations/adoptions/dfsp200t.pdf.



North Carolina has taken approach 3, **broadening guidance** explaining that businesses who have control over virtual currency on behalf of customers are money transmitters.[11]

Texas,[12] **Kansas**,[13] and **Tennessee**[14] have taken approach 2, **narrowing guidance**, explaining that only virtual currency businesses who also deal in traditional currencies (*e.g.* a virtual-currency-for-dollars exchange) are money transmitters. **Illinois** is soliciting comments on proposed guidance that would take this approach as well.[15]

The remaining states have taken approach 1, **do nothing**.

B. Which Approach to Choose?

While Coin Center advocates for a light touch regulatory approach to virtual currency technologies, **we do not encourage states to take a do nothing approach**.

Every state except for **Montana** already regulates money transmitters, requiring that such businesses become licensed *before* taking on customers who are residents of the state. The various statutory definitions of money transmission are broad, focused on older payment systems, and difficult to parse with respect to new technologies and business models emerging in the virtual currency space.

Without some clarifying action from lawmakers or regulators, the vague drafting inherent in these state money transmission statutes leaves open the possibility that a virtual currency company with customers in the state will already qualify as a money transmitter under existing laws, and, if operating without a license, will be subject to substantial civil and criminal punishments. This legal uncertainty and looming liability is a real threat to the talented men and women who develop these technologies or start businesses. At the very least, lawmakers and regulators should be clear when it comes to the application (or non-application) of laws that can so easily ruin the lives and livelihoods of our country's most creative and innovative citizens.

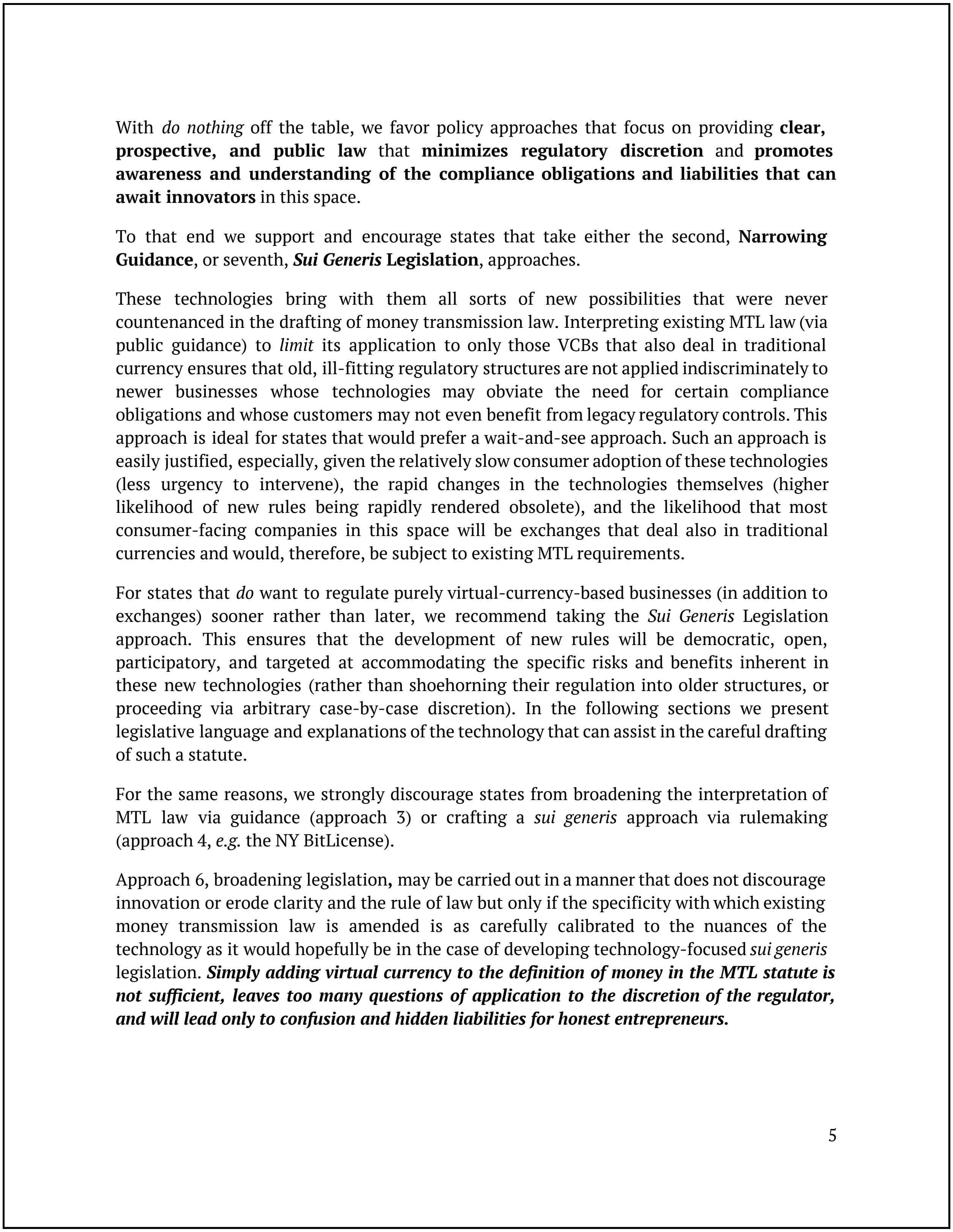
[11] *See* North Carolina Commissioner of Banks, *Money Transmitter Frequently Asked Questions* (last accessed Feb. 2017) http://www.nccob.gov/Public/financialinstitutions/mt/mtfaq.aspx.

[12] *See* Texas Department of Banking, *Supervisory Memorandum 1037* (2014) *available at* http://www.dob.texas.gov/public/uploads/files/consumer-information/sm1037.pdf.

[13] *See* Kansas Office of the State Bank Commissioner, *Guidance Document MT 2014-01* (2014) *available at* http://www.osbckansas.org/mt/guidance/mt2014_01_virtual_currency.pdf

[14] *See* Tennessee Department of Financial Institutions, *Memo: Regulatory Treatment of Virtual Currencies under the Tennessee Money Transmitter Act* (Dec. 2015) http://tn.gov/assets/entities/tdfi/attachments/2015-12-16_TDFI_Memo_on_Virtual_Currency.pdf.

[15] *See* Illinois Department of Financial and Professional Regulation, *Digital Currency Regulatory Guidance* (2017) *available at* https://www.idfpr.com/news/PDFs/IDFPRRequestforCommentsDigitalCurrencyRegulatoryGuidance2016.pdf.



With *do nothing* off the table, we favor policy approaches that focus on providing **clear, prospective, and public law** that **minimizes regulatory discretion** and **promotes awareness and understanding of the compliance obligations and liabilities that can await innovators** in this space.

To that end we support and encourage states that take either the second, **Narrowing Guidance**, or seventh, ***Sui Generis* Legislation**, approaches.

These technologies bring with them all sorts of new possibilities that were never countenanced in the drafting of money transmission law. Interpreting existing MTL law (via public guidance) to *limit* its application to only those VCBs that also deal in traditional currency ensures that old, ill-fitting regulatory structures are not applied indiscriminately to newer businesses whose technologies may obviate the need for certain compliance obligations and whose customers may not even benefit from legacy regulatory controls. This approach is ideal for states that would prefer a wait-and-see approach. Such an approach is easily justified, especially, given the relatively slow consumer adoption of these technologies (less urgency to intervene), the rapid changes in the technologies themselves (higher likelihood of new rules being rapidly rendered obsolete), and the likelihood that most consumer-facing companies in this space will be exchanges that deal also in traditional currencies and would, therefore, be subject to existing MTL requirements.

For states that *do* want to regulate purely virtual-currency-based businesses (in addition to exchanges) sooner rather than later, we recommend taking the *Sui Generis* Legislation approach. This ensures that the development of new rules will be democratic, open, participatory, and targeted at accommodating the specific risks and benefits inherent in these new technologies (rather than shoehorning their regulation into older structures, or proceeding via arbitrary case-by-case discretion). In the following sections we present legislative language and explanations of the technology that can assist in the careful drafting of such a statute.

For the same reasons, we strongly discourage states from broadening the interpretation of MTL law via guidance (approach 3) or crafting a *sui generis* approach via rulemaking (approach 4, *e.g.* the NY BitLicense).

Approach 6, broadening legislation, may be carried out in a manner that does not discourage innovation or erode clarity and the rule of law but only if the specificity with which existing money transmission law is amended is as carefully calibrated to the nuances of the technology as it would hopefully be in the case of developing technology-focused *sui generis* legislation. ***Simply adding virtual currency to the definition of money in the MTL statute is not sufficient, leaves too many questions of application to the discretion of the regulator, and will lead only to confusion and hidden liabilities for honest entrepreneurs.***

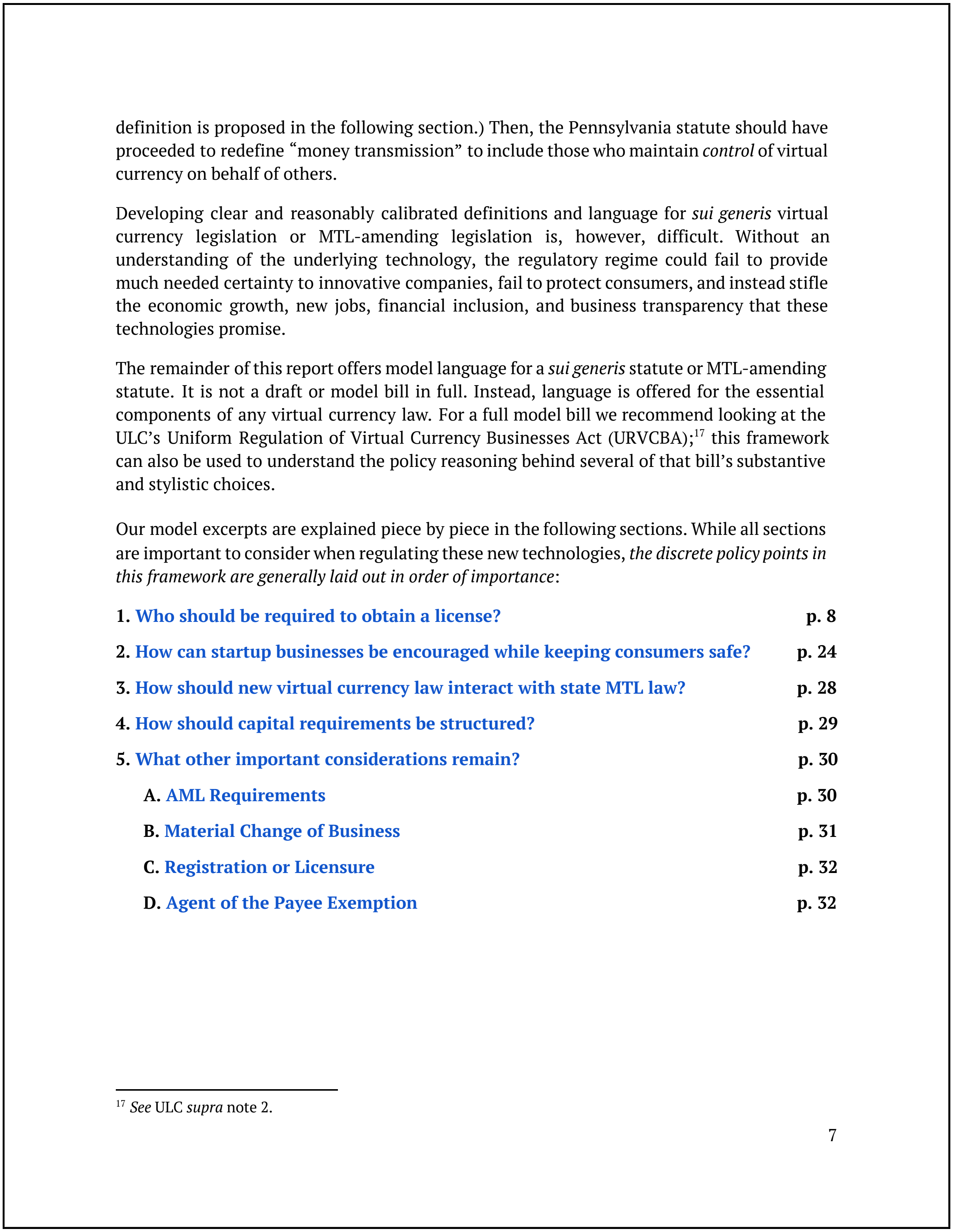
The remainder of this report can also be used as an aid in the process of amending existing money transmission statutes, particularly where simple amendments to existing definitions would result in vague and under- or over-inclusive compliance obligations.

To illustrate, formally re-defining "money" within a statute to include digital or virtual currencies would not be sufficient to guarantee efficient regulation of these new technologies. One must also define what it means to "transmit" a virtual currency or be a "regulated virtual currency transmitter." Traditional money transmission occurs when an intermediary reassigns credits or debits among its customers or partner institutions. These institutions have free reign to assign and reassign credit to different accounts, subject to applicable legal restrictions, as long as they remain solvent at the end of the day. By contrast, bitcoins, for example, can only be transmitted by the holders of unique cryptographic keys. Therefore, only a business that holds these keys could ever have the ability to transmit a bitcoin. A transmittal instrument for a virtual currency is not, then, a promise to pay; it is the ability to pay—*i.e.* cash on hand—as measured by possession or knowledge of cryptographic keys sufficient to execute or prevent a transaction. Just as we would only wish to require licensure from businesses that take it upon themselves to "transmit money" we should only require licensure from VCBs that take it upon themselves to assume ***control*** over keys related to customer bitcoin balances.

For example, a bill was introduced in **Pennsylvania** to amend its money transmission licensing statute in an attempt to cover VCBs.[16] That bill has since failed to pass into law. In an early draft, however, "virtual currency" was added to the definition of "money." The definition of "transmittal instrument" was amended to include "electronic transfer . . . for the payment of money." "Electronic transfer," however, was not defined. Had this draft bill passed in that form, we could reasonably expect a dispute to arise and a judge to interpret the definition in a reasonable manner; however, it seems inefficient to leave such an important distinction to an *ex post* judicial or administrative process. All sorts of individuals and businesses transmit and retransmit Bitcoin transaction messages across the virtual currency's computer network, including individuals running software on their home computers, Internet Service Providers (ISPs), and so-called Bitcoin miners. Unlike the consumer facing bitcoin exchanges who presumably should be licensed, none of these non-controlling entities can spend the bitcoins owned by other participants on the Bitcoin network. However, by playing an infrastructure role in these systems do they take part in an otherwise undefined "electronic transfer?" Why leave this question unresolved in vague legislation and simply hope for a good outcome to follow later in administrative rulings or court cases?

Instead, Pennsylvania should have clearly defined the activity that generates consumer risk: the moment when a VCB is actually has "control" over customer virtual currencies. (Such a

[16] *See* Pennsylvania House Bill 850 (2015) *available at* http://www.legis.state.pa.us/CFDOCS/Legis/PN/Public/btCheck.cfm?txtType=PDF&sessYr=2015&sessInd=0&billBody=H&billTyp=B&billNbr=0850&pn=1029.



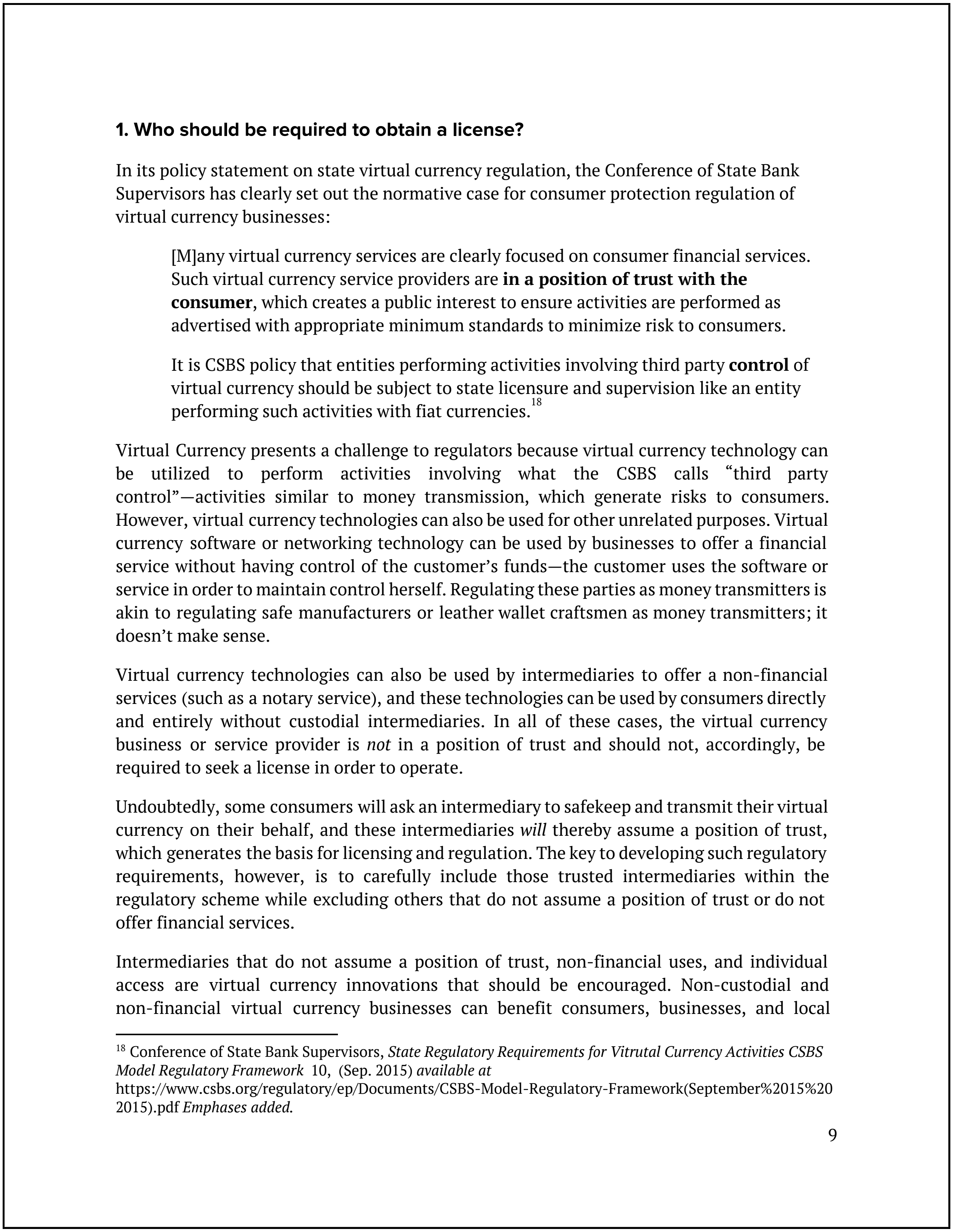
definition is proposed in the following section.) Then, the Pennsylvania statute should have proceeded to redefine “money transmission” to include those who maintain *control* of virtual currency on behalf of others.

Developing clear and reasonably calibrated definitions and language for *sui generis* virtual currency legislation or MTL-amending legislation is, however, difficult. Without an understanding of the underlying technology, the regulatory regime could fail to provide much needed certainty to innovative companies, fail to protect consumers, and instead stifle the economic growth, new jobs, financial inclusion, and business transparency that these technologies promise.

The remainder of this report offers model language for a *sui generis* statute or MTL-amending statute. It is not a draft or model bill in full. Instead, language is offered for the essential components of any virtual currency law. For a full model bill we recommend looking at the ULC’s Uniform Regulation of Virtual Currency Businesses Act (URVCBA);[17] this framework can also be used to understand the policy reasoning behind several of that bill’s substantive and stylistic choices.

Our model excerpts are explained piece by piece in the following sections. While all sections are important to consider when regulating these new technologies, *the discrete policy points in this framework are generally laid out in order of importance*:

[17] *See* ULC *supra* note 2.



1. Who should be required to obtain a license?

In its policy statement on state virtual currency regulation, the Conference of State Bank Supervisors has clearly set out the normative case for consumer protection regulation of virtual currency businesses:

> [M]any virtual currency services are clearly focused on consumer financial services. Such virtual currency service providers are **in a position of trust with the consumer**, which creates a public interest to ensure activities are performed as advertised with appropriate minimum standards to minimize risk to consumers.
>
> It is CSBS policy that entities performing activities involving third party **control** of virtual currency should be subject to state licensure and supervision like an entity performing such activities with fiat currencies.[18]

Virtual Currency presents a challenge to regulators because virtual currency technology can be utilized to perform activities involving what the CSBS calls "third party control"—activities similar to money transmission, which generate risks to consumers. However, virtual currency technologies can also be used for other unrelated purposes. Virtual currency software or networking technology can be used by businesses to offer a financial service without having control of the customer's funds—the customer uses the software or service in order to maintain control herself. Regulating these parties as money transmitters is akin to regulating safe manufacturers or leather wallet craftsmen as money transmitters; it doesn't make sense.

Virtual currency technologies can also be used by intermediaries to offer a non-financial services (such as a notary service), and these technologies can be used by consumers directly and entirely without custodial intermediaries. In all of these cases, the virtual currency business or service provider is *not* in a position of trust and should not, accordingly, be required to seek a license in order to operate.

Undoubtedly, some consumers will ask an intermediary to safekeep and transmit their virtual currency on their behalf, and these intermediaries *will* thereby assume a position of trust, which generates the basis for licensing and regulation. The key to developing such regulatory requirements, however, is to carefully include those trusted intermediaries within the regulatory scheme while excluding others that do not assume a position of trust or do not offer financial services.

Intermediaries that do not assume a position of trust, non-financial uses, and individual access are virtual currency innovations that should be encouraged. Non-custodial and non-financial virtual currency businesses can benefit consumers, businesses, and local

[18] Conference of State Bank Supervisors, *State Regulatory Requirements for Vitrutal Currency Activities CSBS Model Regulatory Framework* 10, (Sep. 2015) *available at* https://www.csbs.org/regulatory/ep/Documents/CSBS-Model-Regulatory-Framework(September%2015%20 2015).pdf *Emphases added.*

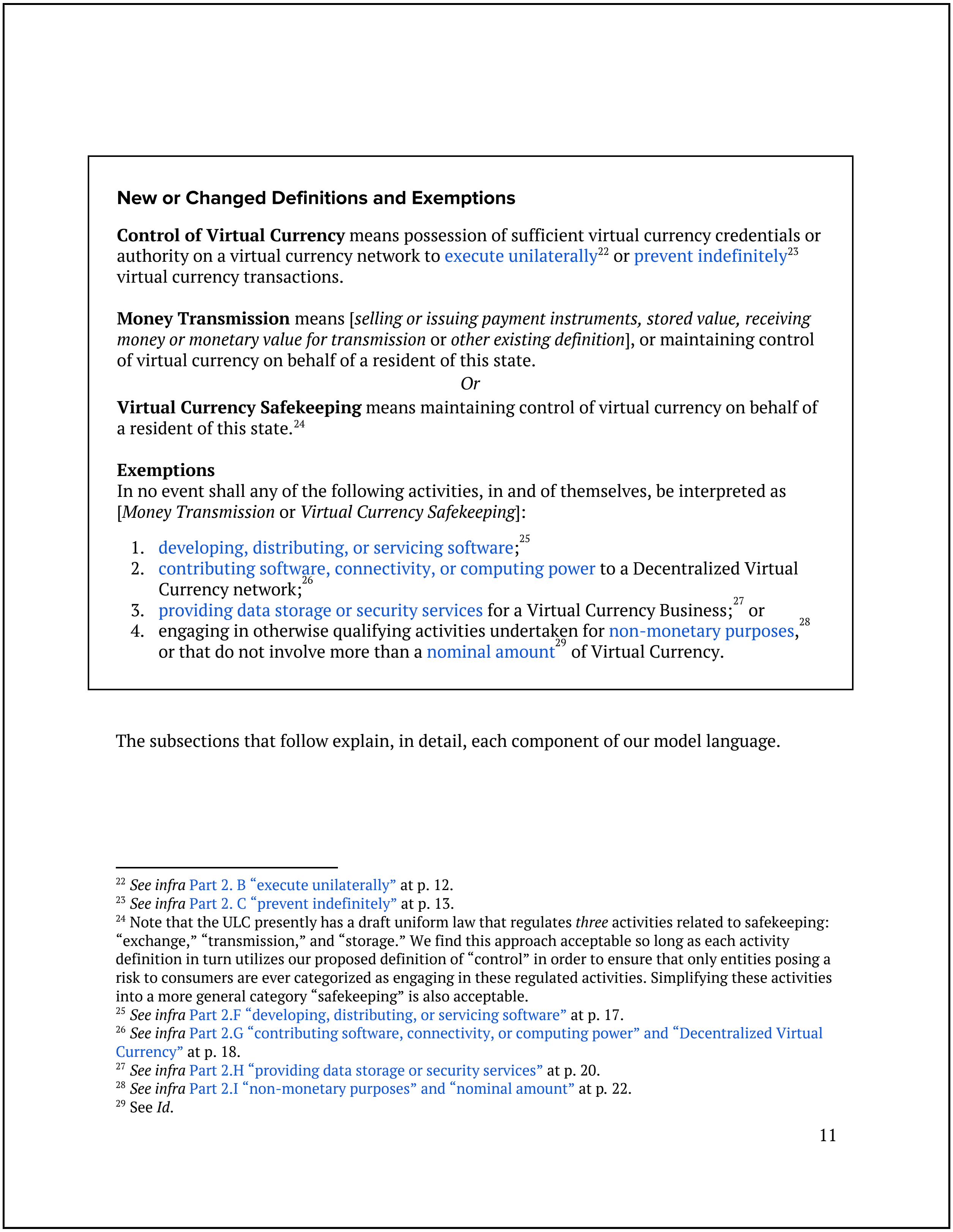
economies through improved financial privacy,[19] financial inclusion,[20] and vibrant technology-based economies. These businesses (as well as any academics or hobbyists experimenting with these technologies) should not be burdened by compliance costs that lack concomitant consumer protection benefits. Neither should these low-risk innovators be in perpetual jeopardy of severe criminal liability for failure to license as a money transmitter. [21]

Custodial intermediaries, on the other hand, so long as they walk and quack like a money transmitting duck, offer the same case for regulation as traditional financial services. The key is narrowly defining that duck. As we will explain in detail, statutes should (1) carefully define "control" of virtual currency as the *de facto* state of holding a customer's virtual currency, (2) use the defined term "control" in a definition of the set of activities that trigger a licensing requirement (*e.g.*, "money transmission" for MTL amendments or "virtual currency safekeeping" for *sui generis* statutes), and then (3) clearly exempt those persons or businesses that do not pose a substantial consumer risk. On the following page is proposed language for these purposes.

[19] *See* Peter Van Valkenburgh, *Bitcoin: Our Best Tool for Privacy and Identity on the Internet*, Coin Center (Mar. 2015) available at https://coincenter.org/2015/03/bitcoin-our-best-tool-for-privacy-and-identity/

[20] *See* Brock Cusick, *How can Bitcoin be Used for Remittances? A Backgrounder for Policymakers*, Coin Center (Dec. 2014) available at https://coincenter.org/2014/12/remittances/.

[21] 18 U.S.C. §1960(a) ("Whoever knowingly conducts, controls, manages, supervises, directs, or owns all or part of an unlicensed money transmitting business, shall be fined in accordance with this title or imprisoned not more than 5 years, or both").



New or Changed Definitions and Exemptions

Control of Virtual Currency means possession of sufficient virtual currency credentials or authority on a virtual currency network to execute unilaterally[22] or prevent indefinitely[23] virtual currency transactions.

Money Transmission means [*selling or issuing payment instruments, stored value, receiving money or monetary value for transmission* or *other existing definition*], or maintaining control of virtual currency on behalf of a resident of this state.

Or

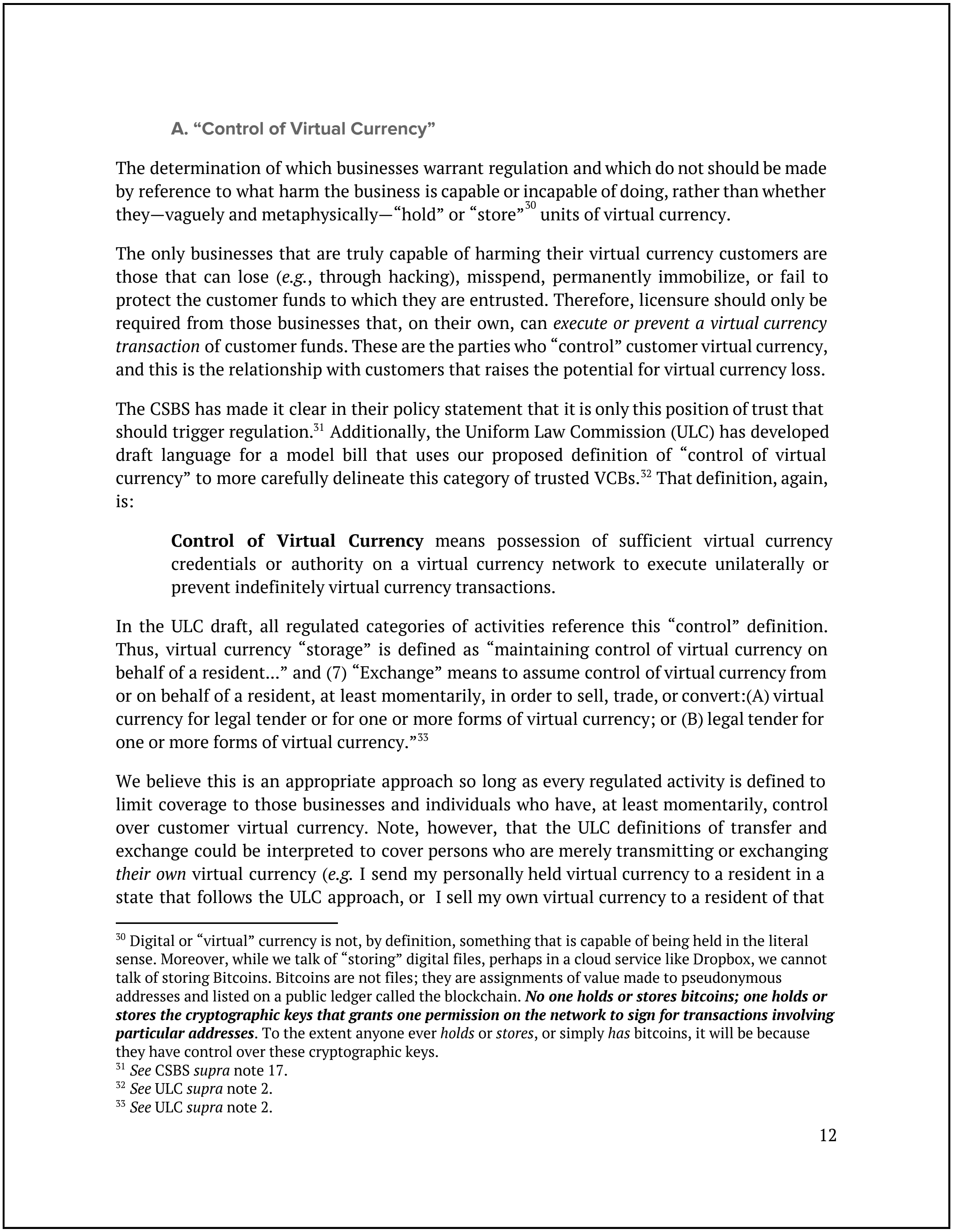
Virtual Currency Safekeeping means maintaining control of virtual currency on behalf of a resident of this state.[24]

Exemptions
In no event shall any of the following activities, in and of themselves, be interpreted as [*Money Transmission* or *Virtual Currency Safekeeping*]:

1. developing, distributing, or servicing software;[25]
2. contributing software, connectivity, or computing power to a Decentralized Virtual Currency network;[26]
3. providing data storage or security services for a Virtual Currency Business;[27] or
4. engaging in otherwise qualifying activities undertaken for non-monetary purposes,[28] or that do not involve more than a nominal amount[29] of Virtual Currency.

The subsections that follow explain, in detail, each component of our model language.

[22] *See infra* Part 2. B "execute unilaterally" at p. 12.
[23] *See infra* Part 2. C "prevent indefinitely" at p. 13.
[24] Note that the ULC presently has a draft uniform law that regulates *three* activities related to safekeeping: "exchange," "transmission," and "storage." We find this approach acceptable so long as each activity definition in turn utilizes our proposed definition of "control" in order to ensure that only entities posing a risk to consumers are ever categorized as engaging in these regulated activities. Simplifying these activities into a more general category "safekeeping" is also acceptable.
[25] *See infra* Part 2.F "developing, distributing, or servicing software" at p. 17.
[26] *See infra* Part 2.G "contributing software, connectivity, or computing power" and "Decentralized Virtual Currency" at p. 18.
[27] *See infra* Part 2.H "providing data storage or security services" at p. 20.
[28] *See infra* Part 2.I "non-monetary purposes" and "nominal amount" at p. 22.
[29] See *Id.*



A. "Control of Virtual Currency"

The determination of which businesses warrant regulation and which do not should be made by reference to what harm the business is capable or incapable of doing, rather than whether they—vaguely and metaphysically—"hold" or "store"[30] units of virtual currency.

The only businesses that are truly capable of harming their virtual currency customers are those that can lose (*e.g.*, through hacking), misspend, permanently immobilize, or fail to protect the customer funds to which they are entrusted. Therefore, licensure should only be required from those businesses that, on their own, can *execute or prevent a virtual currency transaction* of customer funds. These are the parties who "control" customer virtual currency, and this is the relationship with customers that raises the potential for virtual currency loss.

The CSBS has made it clear in their policy statement that it is only this position of trust that should trigger regulation.[31] Additionally, the Uniform Law Commission (ULC) has developed draft language for a model bill that uses our proposed definition of "control of virtual currency" to more carefully delineate this category of trusted VCBs.[32] That definition, again, is:

> **Control of Virtual Currency** means possession of sufficient virtual currency credentials or authority on a virtual currency network to execute unilaterally or prevent indefinitely virtual currency transactions.

In the ULC draft, all regulated categories of activities reference this "control" definition. Thus, virtual currency "storage" is defined as "maintaining control of virtual currency on behalf of a resident…" and (7) "Exchange" means to assume control of virtual currency from or on behalf of a resident, at least momentarily, in order to sell, trade, or convert:(A) virtual currency for legal tender or for one or more forms of virtual currency; or (B) legal tender for one or more forms of virtual currency."[33]

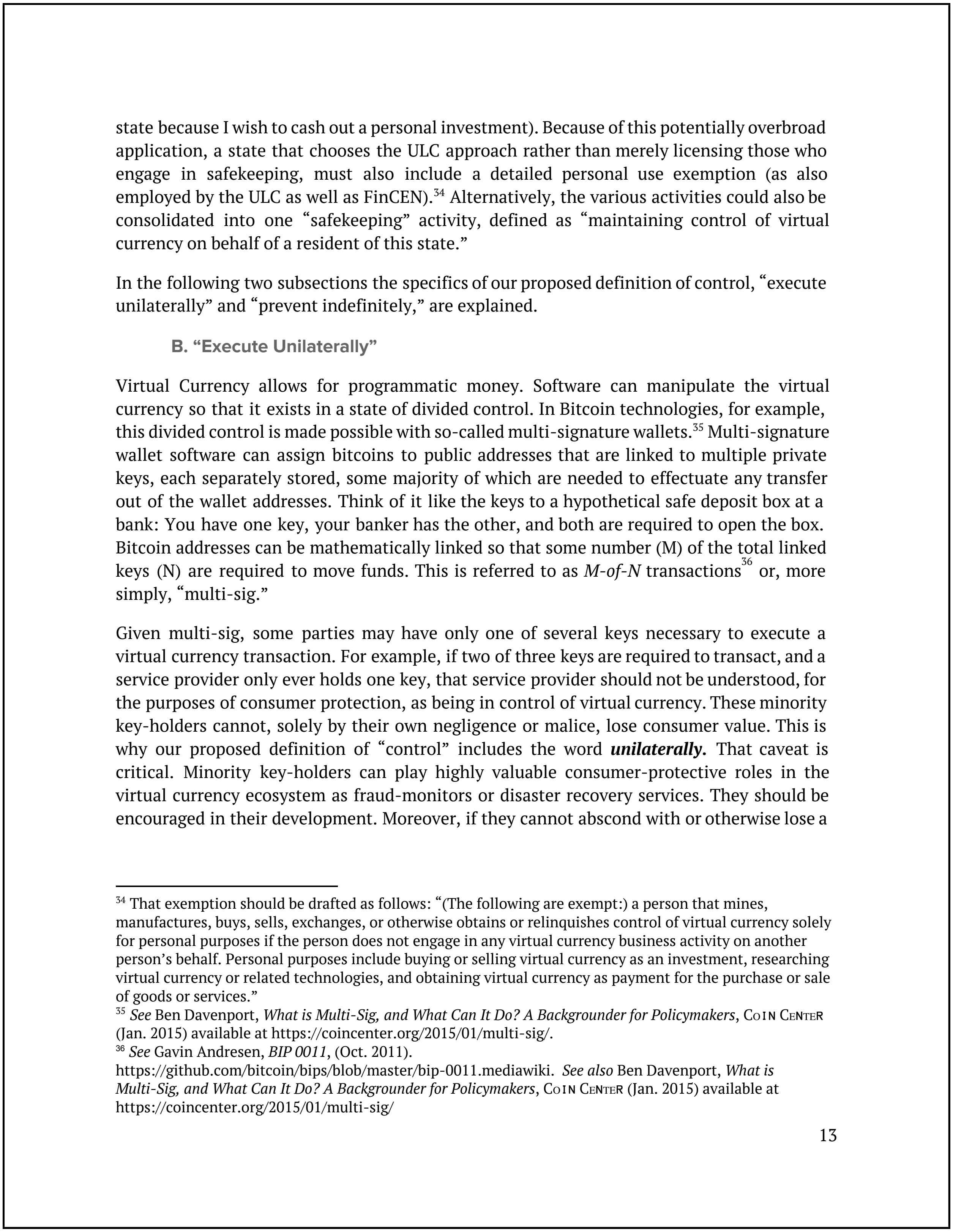
We believe this is an appropriate approach so long as every regulated activity is defined to limit coverage to those businesses and individuals who have, at least momentarily, control over customer virtual currency. Note, however, that the ULC definitions of transfer and exchange could be interpreted to cover persons who are merely transmitting or exchanging *their own* virtual currency (*e.g.* I send my personally held virtual currency to a resident in a state that follows the ULC approach, or I sell my own virtual currency to a resident of that

[30] Digital or "virtual" currency is not, by definition, something that is capable of being held in the literal sense. Moreover, while we talk of "storing" digital files, perhaps in a cloud service like Dropbox, we cannot talk of storing Bitcoins. Bitcoins are not files; they are assignments of value made to pseudonymous addresses and listed on a public ledger called the blockchain. ***No one holds or stores bitcoins; one holds or stores the cryptographic keys that grants one permission on the network to sign for transactions involving particular addresses***. To the extent anyone ever *holds* or *stores*, or simply *has* bitcoins, it will be because they have control over these cryptographic keys.

[31] *See* CSBS *supra* note 17.

[32] *See* ULC *supra* note 2.

[33] *See* ULC *supra* note 2.



state because I wish to cash out a personal investment). Because of this potentially overbroad application, a state that chooses the ULC approach rather than merely licensing those who engage in safekeeping, must also include a detailed personal use exemption (as also employed by the ULC as well as FinCEN).[34] Alternatively, the various activities could also be consolidated into one "safekeeping" activity, defined as "maintaining control of virtual currency on behalf of a resident of this state."

In the following two subsections the specifics of our proposed definition of control, "execute unilaterally" and "prevent indefinitely," are explained.

B. "Execute Unilaterally"

Virtual Currency allows for programmatic money. Software can manipulate the virtual currency so that it exists in a state of divided control. In Bitcoin technologies, for example, this divided control is made possible with so-called multi-signature wallets.[35] Multi-signature wallet software can assign bitcoins to public addresses that are linked to multiple private keys, each separately stored, some majority of which are needed to effectuate any transfer out of the wallet addresses. Think of it like the keys to a hypothetical safe deposit box at a bank: You have one key, your banker has the other, and both are required to open the box. Bitcoin addresses can be mathematically linked so that some number (M) of the total linked keys (N) are required to move funds. This is referred to as *M-of-N* transactions[36] or, more simply, "multi-sig."

Given multi-sig, some parties may have only one of several keys necessary to execute a virtual currency transaction. For example, if two of three keys are required to transact, and a service provider only ever holds one key, that service provider should not be understood, for the purposes of consumer protection, as being in control of virtual currency. These minority key-holders cannot, solely by their own negligence or malice, lose consumer value. This is why our proposed definition of "control" includes the word ***unilaterally.*** That caveat is critical. Minority key-holders can play highly valuable consumer-protective roles in the virtual currency ecosystem as fraud-monitors or disaster recovery services. They should be encouraged in their development. Moreover, if they cannot abscond with or otherwise lose a

[34] That exemption should be drafted as follows: "(The following are exempt:) a person that mines, manufactures, buys, sells, exchanges, or otherwise obtains or relinquishes control of virtual currency solely for personal purposes if the person does not engage in any virtual currency business activity on another person's behalf. Personal purposes include buying or selling virtual currency as an investment, researching virtual currency or related technologies, and obtaining virtual currency as payment for the purchase or sale of goods or services."

[35] *See* Ben Davenport, *What is Multi-Sig, and What Can It Do? A Backgrounder for Policymakers*, Coin Center (Jan. 2015) available at https://coincenter.org/2015/01/multi-sig/.

[36] *See* Gavin Andresen, *BIP 0011*, (Oct. 2011). https://github.com/bitcoin/bips/blob/master/bip-0011.mediawiki. *See also* Ben Davenport, *What is Multi-Sig, and What Can It Do? A Backgrounder for Policymakers*, Coin Center (Jan. 2015) available at https://coincenter.org/2015/01/multi-sig/

customer's funds, mandating their licensure serves no consumer-protective or prudential-regulatory purpose because no solvency risk exists within their business model.

A company could, for example, help store only the disaster recovery key of a customer who is afraid of losing one of her keys or is afraid of her virtual currency exchange (a separate company) being compromised. Another company could, for example, hold a single key to sign off on transactions initiated using the consumer's key after, and only after, the company verifies that the consumer's phone has not been hacked or her key otherwise compromised.

Both of these hypothetical companies would provide an essential service in securing and safeguarding customer funds. Both hypothetical services are novel and unavailable to the customers of traditional banks and money transmitters because they rely on the use of new cryptographic tools and the blockchain to divide control among multiple businesses without using laws to enforce that division. Neither of these companies, however, should need to be licensed as money transmitters. Without possession of *sufficient* keys to move or immobilize a customer's funds on its own, the company does not pose a consumer protection risk; quite the opposite, they mitigate that risk.

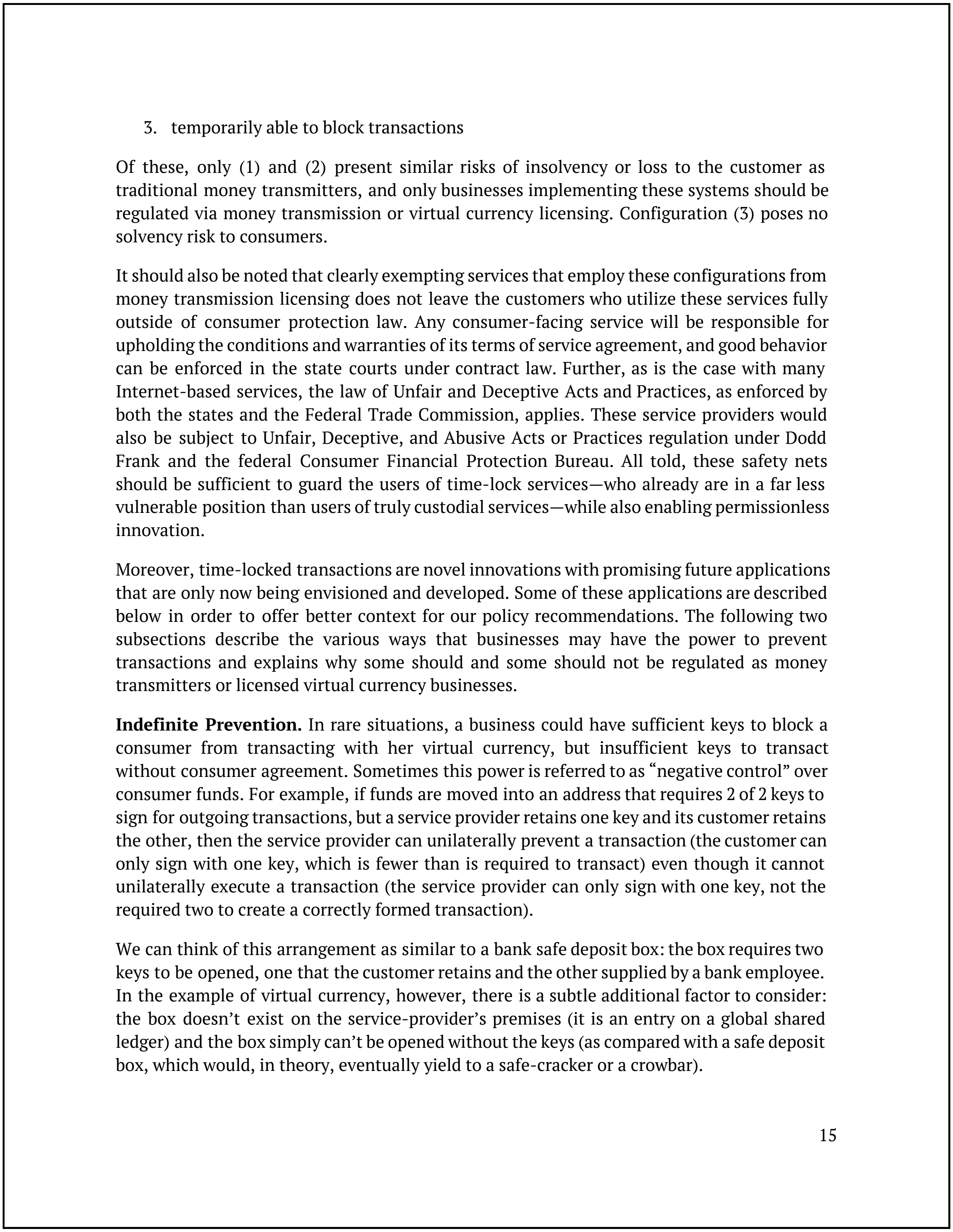
Such companies will be highly valuable innovators in the field of virtual currency. The technology that enables divided key control, *i.e.*, multi-sig, is widely understood within the industry as the single best tool for preventing hacking thefts.[37] Defining control to include only those who can **unilaterally** execute a transaction, ensures that these tools can be developed without subjecting their creators to a licensing regime that adds costs to their business without delivering any benefits to consumers. This definition also sends a credible and welcome signal to innovators in the virtual currency space: *we value your effort to build technology that will complement our consumer protection efforts and do not want to impede your progress unnecessarily*.

C. "Prevent Indefinitely"

Given multi-sig, we can imagine many various business models where control of funds is divided between the customer and the business or even between a customer and multiple businesses. Additionally, useful systems can be designed by using another technology native to many cryptocurrencies: time-locked transactions (within the Bitcoin protocol referred to as n-lock transactions). With time-locked transactions a business may temporarily have the ability to stop a user from transacting with some certain amount of cryptocurrency, but the user will always automatically regain full control of the funds after a specified time. So, in full, our list of possible configurations for service-providers is as follows:

1. unilaterally able to transact on user's behalf
2. able to block transaction on user's behalf indefinitely

[37] *See* Ben Davenport, *No Sleep Till Multi-Sig* (Jan. 12, 2015) https://medium.com/@bendavenport/no-sleep-till-multi-sig-7db367998bc7.



3. temporarily able to block transactions

Of these, only (1) and (2) present similar risks of insolvency or loss to the customer as traditional money transmitters, and only businesses implementing these systems should be regulated via money transmission or virtual currency licensing. Configuration (3) poses no solvency risk to consumers.

It should also be noted that clearly exempting services that employ these configurations from money transmission licensing does not leave the customers who utilize these services fully outside of consumer protection law. Any consumer-facing service will be responsible for upholding the conditions and warranties of its terms of service agreement, and good behavior can be enforced in the state courts under contract law. Further, as is the case with many Internet-based services, the law of Unfair and Deceptive Acts and Practices, as enforced by both the states and the Federal Trade Commission, applies. These service providers would also be subject to Unfair, Deceptive, and Abusive Acts or Practices regulation under Dodd Frank and the federal Consumer Financial Protection Bureau. All told, these safety nets should be sufficient to guard the users of time-lock services—who already are in a far less vulnerable position than users of truly custodial services—while also enabling permissionless innovation.

Moreover, time-locked transactions are novel innovations with promising future applications that are only now being envisioned and developed. Some of these applications are described below in order to offer better context for our policy recommendations. The following two subsections describe the various ways that businesses may have the power to prevent transactions and explains why some should and some should not be regulated as money transmitters or licensed virtual currency businesses.

Indefinite Prevention. In rare situations, a business could have sufficient keys to block a consumer from transacting with her virtual currency, but insufficient keys to transact without consumer agreement. Sometimes this power is referred to as "negative control" over consumer funds. For example, if funds are moved into an address that requires 2 of 2 keys to sign for outgoing transactions, but a service provider retains one key and its customer retains the other, then the service provider can unilaterally prevent a transaction (the customer can only sign with one key, which is fewer than is required to transact) even though it cannot unilaterally execute a transaction (the service provider can only sign with one key, not the required two to create a correctly formed transaction).

We can think of this arrangement as similar to a bank safe deposit box: the box requires two keys to be opened, one that the customer retains and the other supplied by a bank employee. In the example of virtual currency, however, there is a subtle additional factor to consider: the box doesn't exist on the service-provider's premises (it is an entry on a global shared ledger) and the box simply can't be opened without the keys (as compared with a safe deposit box, which would, in theory, eventually yield to a safe-cracker or a crowbar).

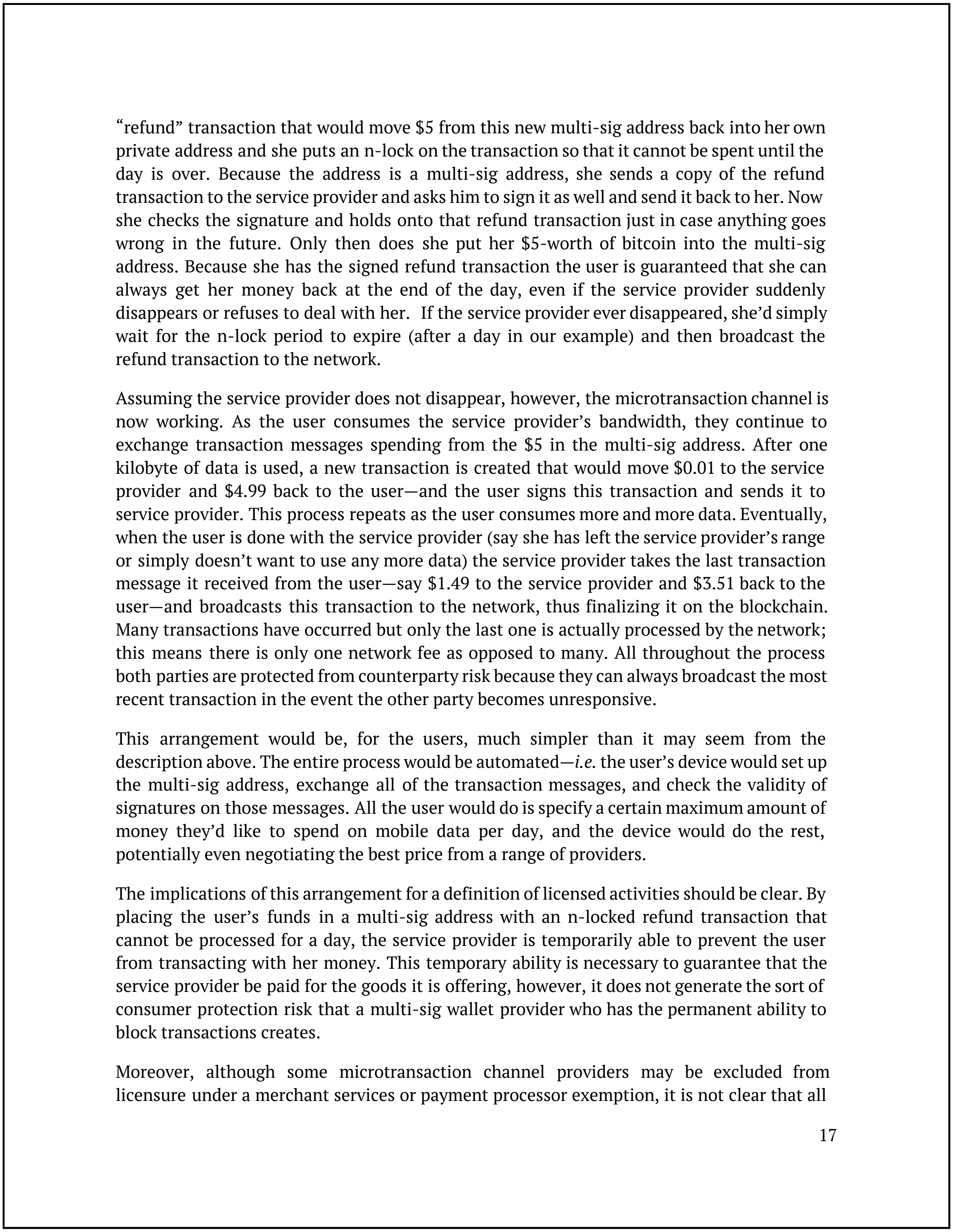
It is unclear why a business would ever set up such an arrangement. However, if it does so it should be regulated as any other money transmitter. Should the business ever be hacked, for example, the hackers could take the key and blackmail the consumer into signing with the other key for a transaction that would send some funds to the thieves' address and some to another address held by the customer. The blackmailers will probably succeed in this scam, given that refusal to comply will irrevocably lock all of the funds out of anyone's reach. Because of this vulnerability, businesses unable to execute but able to indefinitely prevent transactions pose similar risks to consumers and assume a similar level of trust as traditional money transmitters. They should be regulated accordingly.

Temporary Prevention. In the most fundamental sense, the transaction validators—*e.g.* miners in the case of Bitcoin—on a cryptocurrency network will be capable of preventing transactions for the brief period (typically around 10 minutes or less) in which they are capable of incorporating or not incorporating requested transactions into the currency's blockchain. Additionally, as discussed, the Bitcoin protocol also allows for transactions that are time-locked—often referred to as "n-lock" transactions. An n-lock transaction can be signed by the party moving funds but in such a way that it cannot be accepted by the network until a specified time in the future.

A primary use for n-lock transactions is in the creation of low-trust microtransaction channels for the metering of goods or services.[38] Say, for example, you were a cellular network provider and you wanted to charge your network users for every kilobyte of data they used. Rather than establishing a legal relationship with the user—*e.g.* signing them up for a subscription or otherwise making a formal service contract—you'd like to allow anyone to connect to your network, sight unseen, and have their phone automatically pay you for its data usage. Writing a new microtransaction to the blockchain for every kilobyte of data consumed is not an efficient method to create such a system. Even Bitcoin—often celebrated for its low per-transaction fees relative to credit card networks—would require some fees for each transaction, and if an additional transaction was required for every few seconds or minutes of additional use, the cumulative fees would still be cost-prohibitive. Bitcoin, and other cryptocurrencies, however, can use n-lock transactions and microtransaction channels to achieve the same result with extremely low fees.

To set up a microtransaction channel the user's device and the service provider's server generate a new 2-of-2 multi-sig address. The user retains one key and the service provider gets the other. Into this address the user will put the maximum amount of bitcoin she imagines spending on mobile data with this provider over a set period. Let's say $5 for the day. Before moving any of her funds into this multi-sig address, however, the user writes a

[38] For a more complete backgrounder on microtransaction channels see Chris Smith, "What are Micropayments and How does Bitcoin Enable Them? A Backgrounder for Policymakers," *Coin Center* (June 2015) available at https://coincenter.org/2015/06/what-are-micropayments-and-how-does-bitcoin-enable-them/.



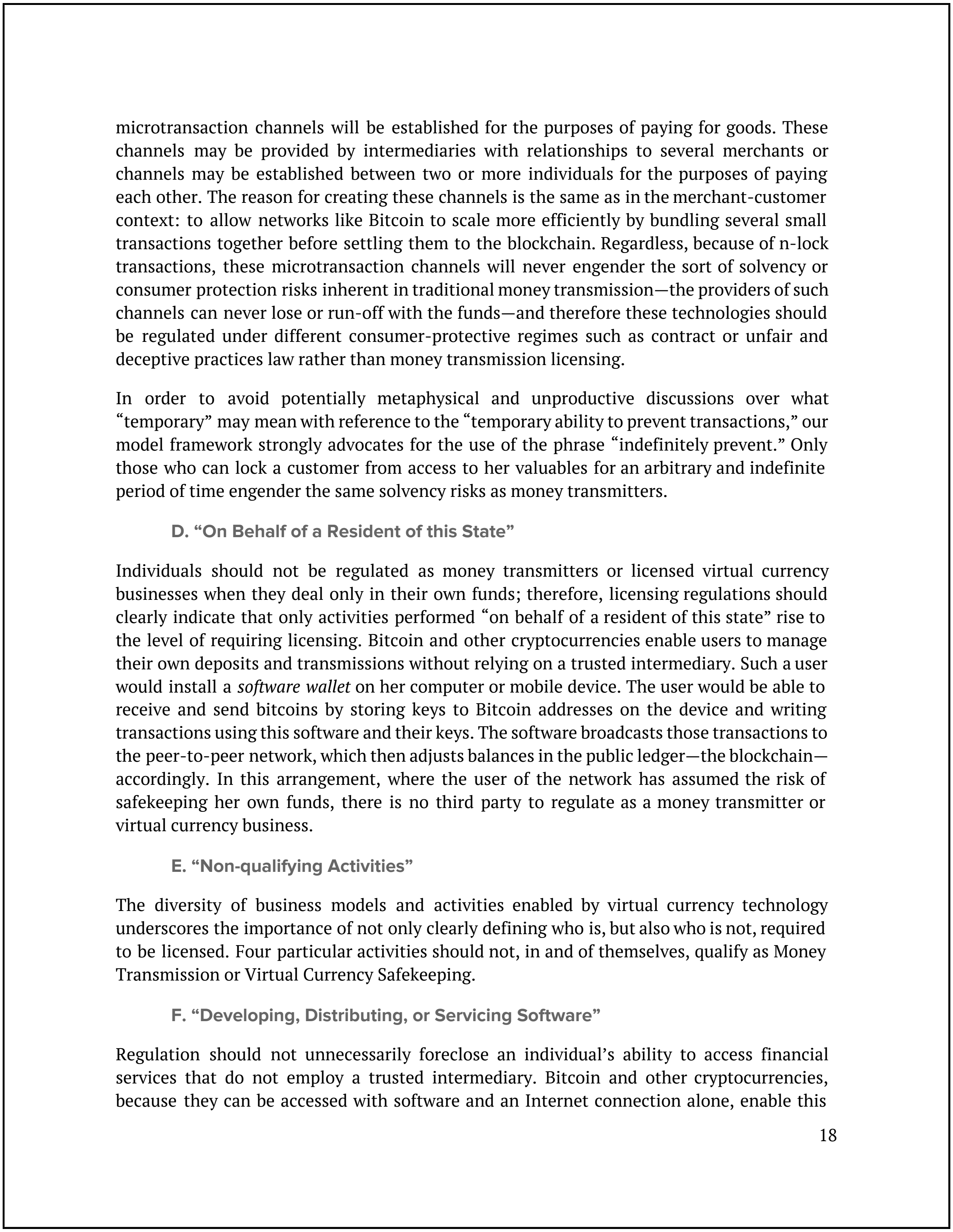
"refund" transaction that would move $5 from this new multi-sig address back into her own private address and she puts an n-lock on the transaction so that it cannot be spent until the day is over. Because the address is a multi-sig address, she sends a copy of the refund transaction to the service provider and asks him to sign it as well and send it back to her. Now she checks the signature and holds onto that refund transaction just in case anything goes wrong in the future. Only then does she put her $5-worth of bitcoin into the multi-sig address. Because she has the signed refund transaction the user is guaranteed that she can always get her money back at the end of the day, even if the service provider suddenly disappears or refuses to deal with her. If the service provider ever disappeared, she'd simply wait for the n-lock period to expire (after a day in our example) and then broadcast the refund transaction to the network.

Assuming the service provider does not disappear, however, the microtransaction channel is now working. As the user consumes the service provider's bandwidth, they continue to exchange transaction messages spending from the $5 in the multi-sig address. After one kilobyte of data is used, a new transaction is created that would move $0.01 to the service provider and $4.99 back to the user—and the user signs this transaction and sends it to service provider. This process repeats as the user consumes more and more data. Eventually, when the user is done with the service provider (say she has left the service provider's range or simply doesn't want to use any more data) the service provider takes the last transaction message it received from the user—say $1.49 to the service provider and $3.51 back to the user—and broadcasts this transaction to the network, thus finalizing it on the blockchain. Many transactions have occurred but only the last one is actually processed by the network; this means there is only one network fee as opposed to many. All throughout the process both parties are protected from counterparty risk because they can always broadcast the most recent transaction in the event the other party becomes unresponsive.

This arrangement would be, for the users, much simpler than it may seem from the description above. The entire process would be automated—*i.e.* the user's device would set up the multi-sig address, exchange all of the transaction messages, and check the validity of signatures on those messages. All the user would do is specify a certain maximum amount of money they'd like to spend on mobile data per day, and the device would do the rest, potentially even negotiating the best price from a range of providers.

The implications of this arrangement for a definition of licensed activities should be clear. By placing the user's funds in a multi-sig address with an n-locked refund transaction that cannot be processed for a day, the service provider is temporarily able to prevent the user from transacting with her money. This temporary ability is necessary to guarantee that the service provider be paid for the goods it is offering, however, it does not generate the sort of consumer protection risk that a multi-sig wallet provider who has the permanent ability to block transactions creates.

Moreover, although some microtransaction channel providers may be excluded from licensure under a merchant services or payment processor exemption, it is not clear that all



microtransaction channels will be established for the purposes of paying for goods. These channels may be provided by intermediaries with relationships to several merchants or channels may be established between two or more individuals for the purposes of paying each other. The reason for creating these channels is the same as in the merchant-customer context: to allow networks like Bitcoin to scale more efficiently by bundling several small transactions together before settling them to the blockchain. Regardless, because of n-lock transactions, these microtransaction channels will never engender the sort of solvency or consumer protection risks inherent in traditional money transmission—the providers of such channels can never lose or run-off with the funds—and therefore these technologies should be regulated under different consumer-protective regimes such as contract or unfair and deceptive practices law rather than money transmission licensing.

In order to avoid potentially metaphysical and unproductive discussions over what "temporary" may mean with reference to the "temporary ability to prevent transactions," our model framework strongly advocates for the use of the phrase "indefinitely prevent." Only those who can lock a customer from access to her valuables for an arbitrary and indefinite period of time engender the same solvency risks as money transmitters.

D. "On Behalf of a Resident of this State"

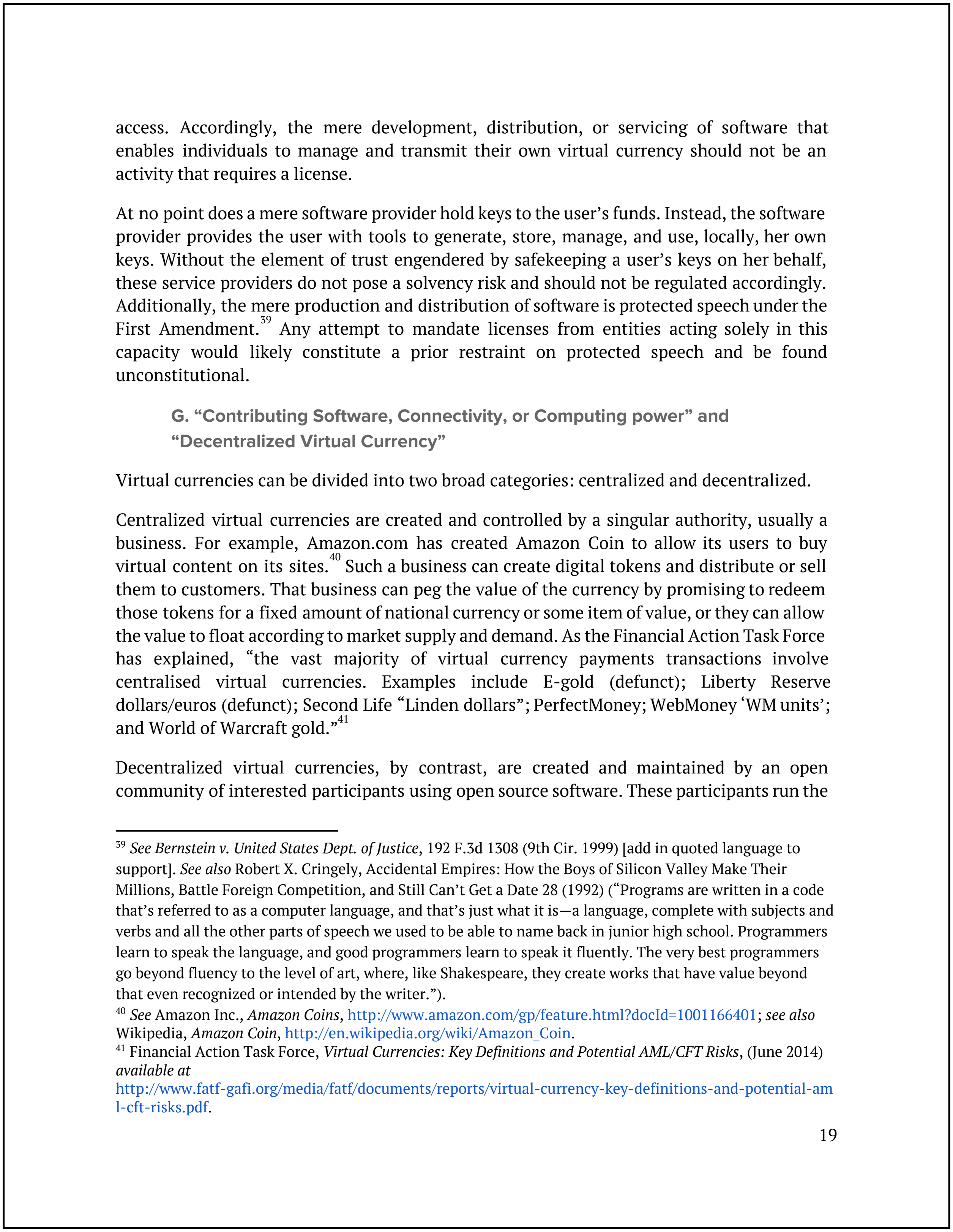
Individuals should not be regulated as money transmitters or licensed virtual currency businesses when they deal only in their own funds; therefore, licensing regulations should clearly indicate that only activities performed "on behalf of a resident of this state" rise to the level of requiring licensing. Bitcoin and other cryptocurrencies enable users to manage their own deposits and transmissions without relying on a trusted intermediary. Such a user would install a *software wallet* on her computer or mobile device. The user would be able to receive and send bitcoins by storing keys to Bitcoin addresses on the device and writing transactions using this software and their keys. The software broadcasts those transactions to the peer-to-peer network, which then adjusts balances in the public ledger—the blockchain—accordingly. In this arrangement, where the user of the network has assumed the risk of safekeeping her own funds, there is no third party to regulate as a money transmitter or virtual currency business.

E. "Non-qualifying Activities"

The diversity of business models and activities enabled by virtual currency technology underscores the importance of not only clearly defining who is, but also who is not, required to be licensed. Four particular activities should not, in and of themselves, qualify as Money Transmission or Virtual Currency Safekeeping.

F. "Developing, Distributing, or Servicing Software"

Regulation should not unnecessarily foreclose an individual's ability to access financial services that do not employ a trusted intermediary. Bitcoin and other cryptocurrencies, because they can be accessed with software and an Internet connection alone, enable this



access. Accordingly, the mere development, distribution, or servicing of software that enables individuals to manage and transmit their own virtual currency should not be an activity that requires a license.

At no point does a mere software provider hold keys to the user's funds. Instead, the software provider provides the user with tools to generate, store, manage, and use, locally, her own keys. Without the element of trust engendered by safekeeping a user's keys on her behalf, these service providers do not pose a solvency risk and should not be regulated accordingly. Additionally, the mere production and distribution of software is protected speech under the First Amendment.[39] Any attempt to mandate licenses from entities acting solely in this capacity would likely constitute a prior restraint on protected speech and be found unconstitutional.

G. "Contributing Software, Connectivity, or Computing power" and "Decentralized Virtual Currency"

Virtual currencies can be divided into two broad categories: centralized and decentralized.

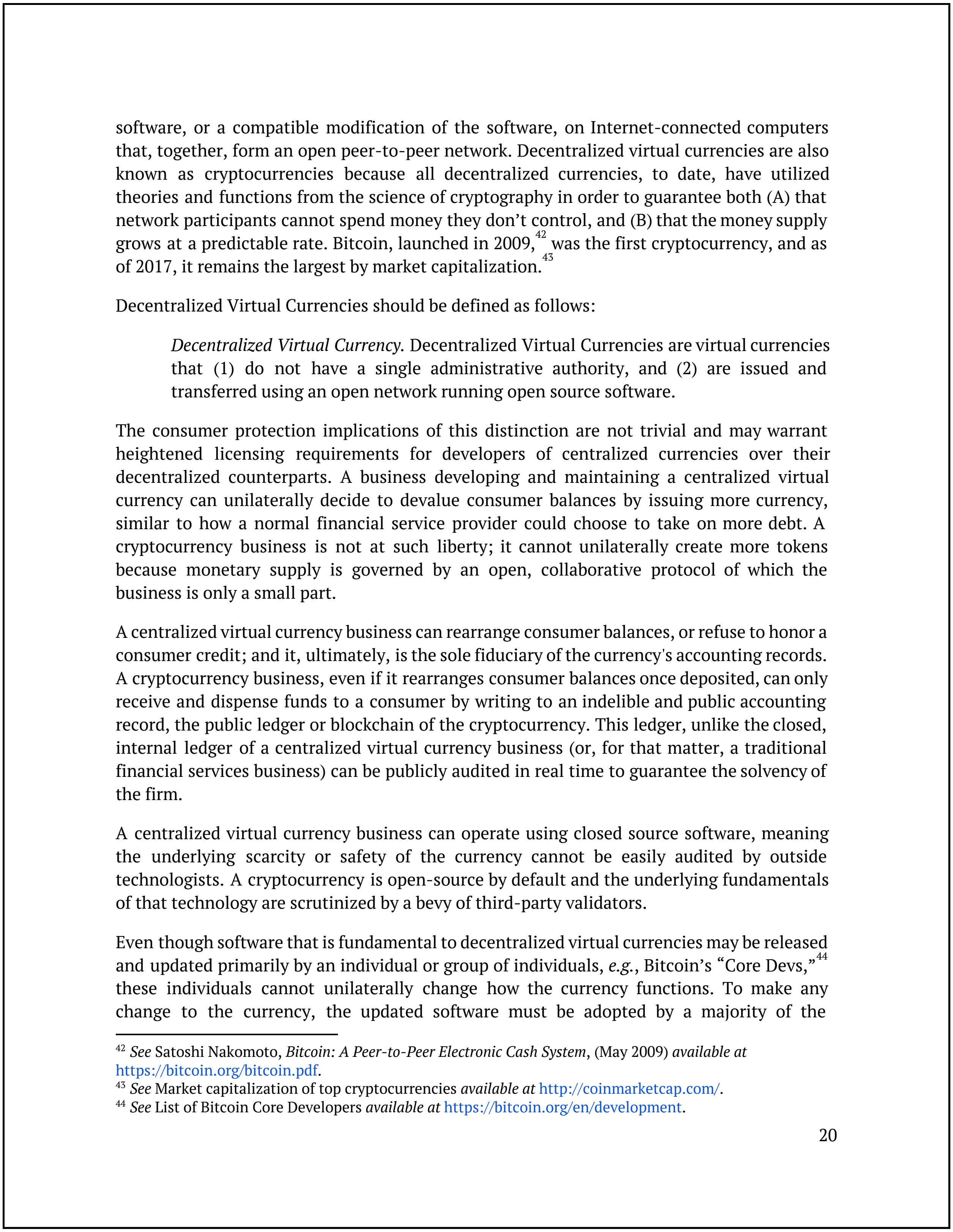
Centralized virtual currencies are created and controlled by a singular authority, usually a business. For example, Amazon.com has created Amazon Coin to allow its users to buy virtual content on its sites.[40] Such a business can create digital tokens and distribute or sell them to customers. That business can peg the value of the currency by promising to redeem those tokens for a fixed amount of national currency or some item of value, or they can allow the value to float according to market supply and demand. As the Financial Action Task Force has explained, "the vast majority of virtual currency payments transactions involve centralised virtual currencies. Examples include E-gold (defunct); Liberty Reserve dollars/euros (defunct); Second Life "Linden dollars"; PerfectMoney; WebMoney 'WM units'; and World of Warcraft gold."[41]

Decentralized virtual currencies, by contrast, are created and maintained by an open community of interested participants using open source software. These participants run the

[39] *See Bernstein v. United States Dept. of Justice*, 192 F.3d 1308 (9th Cir. 1999) [add in quoted language to support]. *See also* Robert X. Cringely, Accidental Empires: How the Boys of Silicon Valley Make Their Millions, Battle Foreign Competition, and Still Can't Get a Date 28 (1992) ("Programs are written in a code that's referred to as a computer language, and that's just what it is—a language, complete with subjects and verbs and all the other parts of speech we used to be able to name back in junior high school. Programmers learn to speak the language, and good programmers learn to speak it fluently. The very best programmers go beyond fluency to the level of art, where, like Shakespeare, they create works that have value beyond that even recognized or intended by the writer.").

[40] *See* Amazon Inc., *Amazon Coins*, http://www.amazon.com/gp/feature.html?docId=1001166401; *see also* Wikipedia, *Amazon Coin*, http://en.wikipedia.org/wiki/Amazon_Coin.

[41] Financial Action Task Force, *Virtual Currencies: Key Definitions and Potential AML/CFT Risks*, (June 2014) *available at* http://www.fatf-gafi.org/media/fatf/documents/reports/virtual-currency-key-definitions-and-potential-aml-cft-risks.pdf.



software, or a compatible modification of the software, on Internet-connected computers that, together, form an open peer-to-peer network. Decentralized virtual currencies are also known as cryptocurrencies because all decentralized currencies, to date, have utilized theories and functions from the science of cryptography in order to guarantee both (A) that network participants cannot spend money they don't control, and (B) that the money supply grows at a predictable rate. Bitcoin, launched in 2009,[42] was the first cryptocurrency, and as of 2017, it remains the largest by market capitalization.[43]

Decentralized Virtual Currencies should be defined as follows:

> *Decentralized Virtual Currency.* Decentralized Virtual Currencies are virtual currencies that (1) do not have a single administrative authority, and (2) are issued and transferred using an open network running open source software.

The consumer protection implications of this distinction are not trivial and may warrant heightened licensing requirements for developers of centralized currencies over their decentralized counterparts. A business developing and maintaining a centralized virtual currency can unilaterally decide to devalue consumer balances by issuing more currency, similar to how a normal financial service provider could choose to take on more debt. A cryptocurrency business is not at such liberty; it cannot unilaterally create more tokens because monetary supply is governed by an open, collaborative protocol of which the business is only a small part.

A centralized virtual currency business can rearrange consumer balances, or refuse to honor a consumer credit; and it, ultimately, is the sole fiduciary of the currency's accounting records. A cryptocurrency business, even if it rearranges consumer balances once deposited, can only receive and dispense funds to a consumer by writing to an indelible and public accounting record, the public ledger or blockchain of the cryptocurrency. This ledger, unlike the closed, internal ledger of a centralized virtual currency business (or, for that matter, a traditional financial services business) can be publicly audited in real time to guarantee the solvency of the firm.

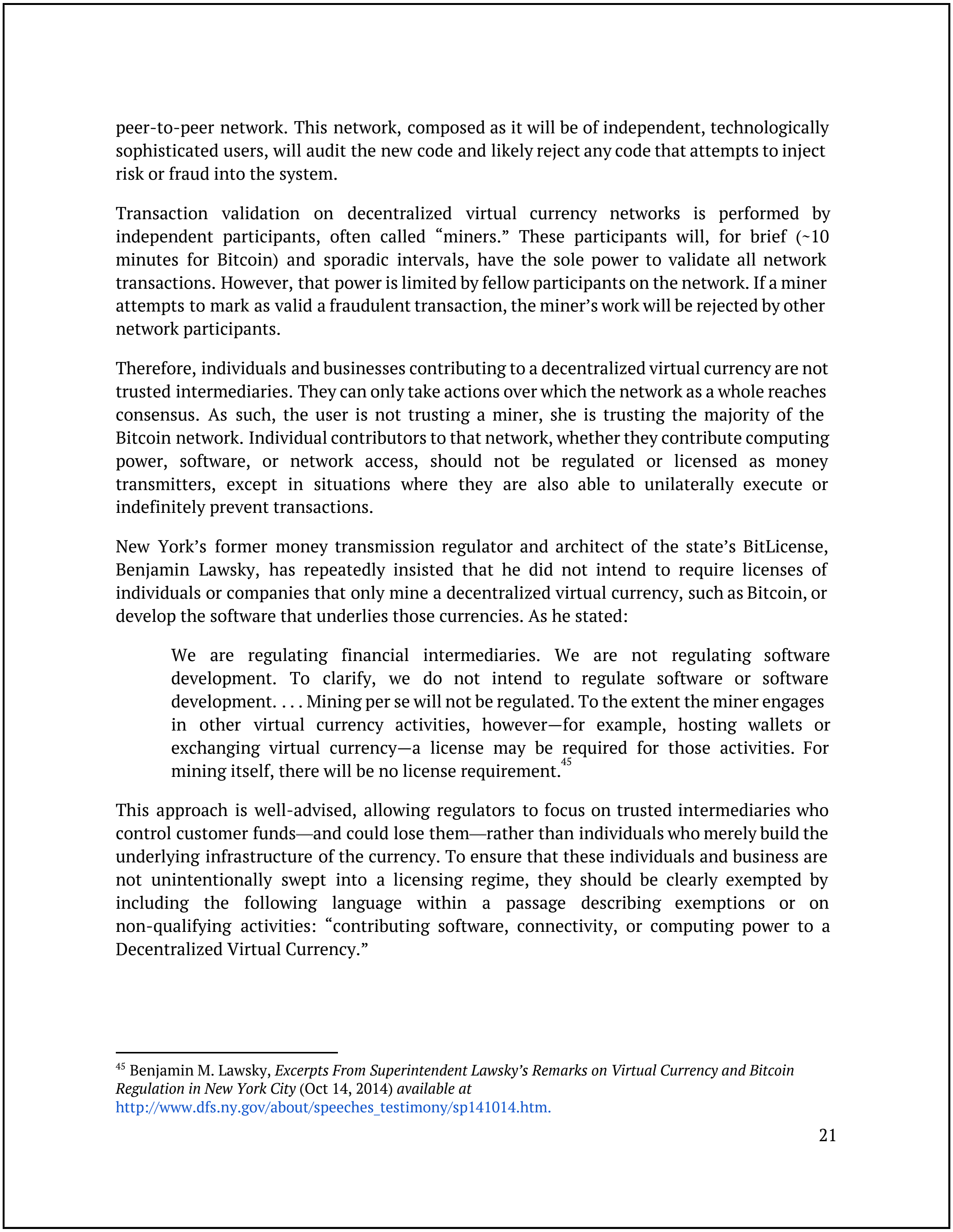
A centralized virtual currency business can operate using closed source software, meaning the underlying scarcity or safety of the currency cannot be easily audited by outside technologists. A cryptocurrency is open-source by default and the underlying fundamentals of that technology are scrutinized by a bevy of third-party validators.

Even though software that is fundamental to decentralized virtual currencies may be released and updated primarily by an individual or group of individuals, *e.g.*, Bitcoin's "Core Devs,"[44] these individuals cannot unilaterally change how the currency functions. To make any change to the currency, the updated software must be adopted by a majority of the

[42] *See* Satoshi Nakomoto, *Bitcoin: A Peer-to-Peer Electronic Cash System*, (May 2009) *available at* https://bitcoin.org/bitcoin.pdf.

[43] *See* Market capitalization of top cryptocurrencies *available at* http://coinmarketcap.com/.

[44] *See* List of Bitcoin Core Developers *available at* https://bitcoin.org/en/development.



peer-to-peer network. This network, composed as it will be of independent, technologically sophisticated users, will audit the new code and likely reject any code that attempts to inject risk or fraud into the system.

Transaction validation on decentralized virtual currency networks is performed by independent participants, often called “miners.” These participants will, for brief (~10 minutes for Bitcoin) and sporadic intervals, have the sole power to validate all network transactions. However, that power is limited by fellow participants on the network. If a miner attempts to mark as valid a fraudulent transaction, the miner’s work will be rejected by other network participants.

Therefore, individuals and businesses contributing to a decentralized virtual currency are not trusted intermediaries. They can only take actions over which the network as a whole reaches consensus. As such, the user is not trusting a miner, she is trusting the majority of the Bitcoin network. Individual contributors to that network, whether they contribute computing power, software, or network access, should not be regulated or licensed as money transmitters, except in situations where they are also able to unilaterally execute or indefinitely prevent transactions.

New York’s former money transmission regulator and architect of the state’s BitLicense, Benjamin Lawsky, has repeatedly insisted that he did not intend to require licenses of individuals or companies that only mine a decentralized virtual currency, such as Bitcoin, or develop the software that underlies those currencies. As he stated:

> We are regulating financial intermediaries. We are not regulating software development. To clarify, we do not intend to regulate software or software development. . . . Mining per se will not be regulated. To the extent the miner engages in other virtual currency activities, however—for example, hosting wallets or exchanging virtual currency—a license may be required for those activities. For mining itself, there will be no license requirement.[45]

This approach is well-advised, allowing regulators to focus on trusted intermediaries who control customer funds—and could lose them—rather than individuals who merely build the underlying infrastructure of the currency. To ensure that these individuals and business are not unintentionally swept into a licensing regime, they should be clearly exempted by including the following language within a passage describing exemptions or on non-qualifying activities: “contributing software, connectivity, or computing power to a Decentralized Virtual Currency.”

[45] Benjamin M. Lawsky, *Excerpts From Superintendent Lawsky’s Remarks on Virtual Currency and Bitcoin Regulation in New York City* (Oct 14, 2014) *available at* http://www.dfs.ny.gov/about/speeches_testimony/sp141014.htm.

H. “Providing Data Storage or Security Services”

As the Bitcoin ecosystem has matured, a new class of infrastructure service providers has emerged. Interacting with the Bitcoin protocol can be technically complex, particularly when using advanced transactions such as the multi-sig or divided key transactions described in a previous section.[46] Early bitcoin hosted wallet providers and exchanges generally coded these transactions in-house. However, this activity may not be the organization’s expertise or comparative advantage. A consumer-facing business may find it more advantageous to focus on marketing, user experience, and regulatory compliance. It may, therefore, choose to contract-out the safekeeping of customer bitcoin keys to business-to-business firms that have developed expertise at utilizing multi-signature transactions and cold storage in order to best secure sensitive data.[47]

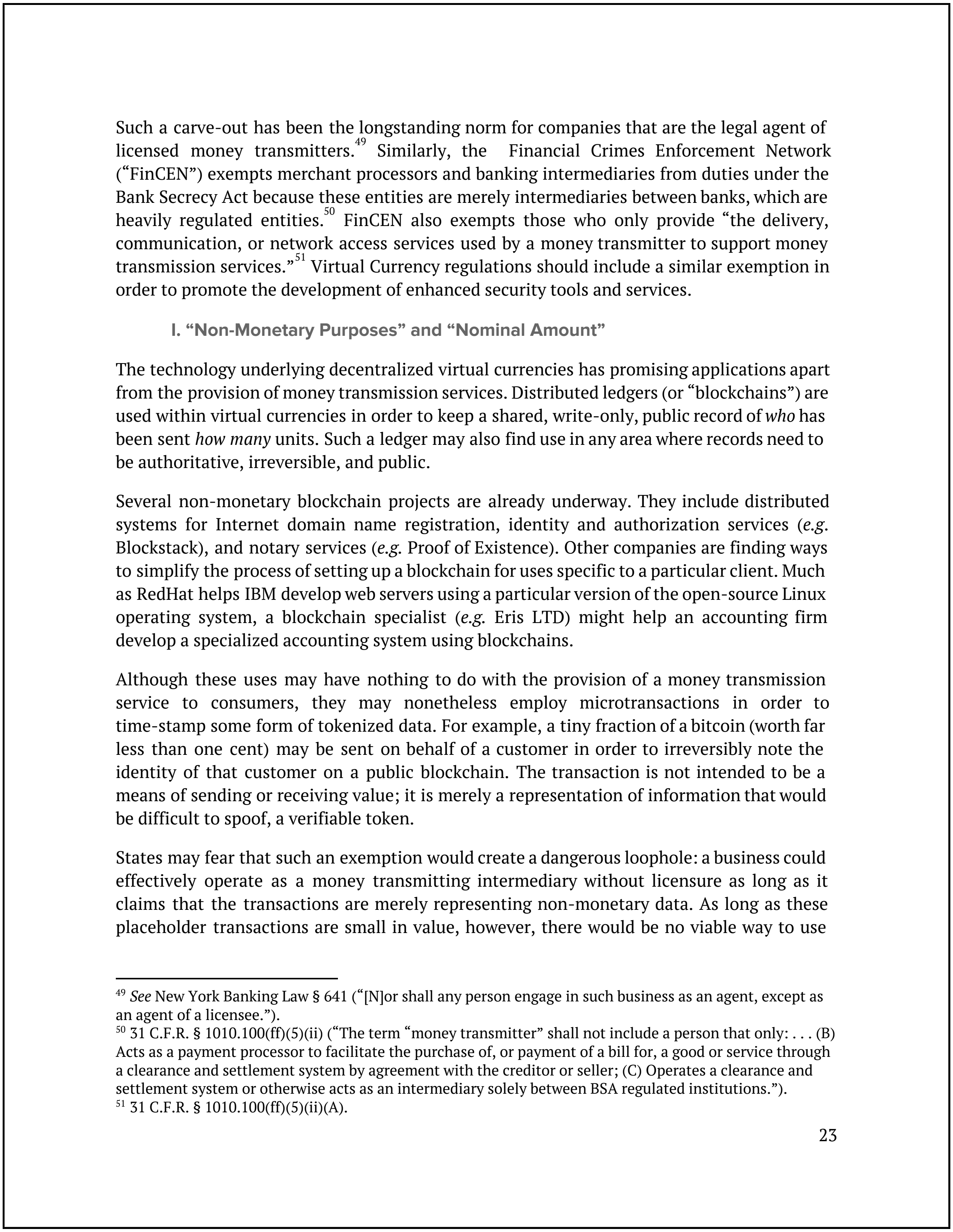
This is not novel in the world of Internet technologies. The video-on-demand service Netflix, for example, does not actually build or maintain the technology necessary to store video data. Instead, it relies on Amazon’s cloud storage solution, Amazon Web Services.[48] If a Bitcoin hosted wallet provider or exchange decided to contract-out the safekeeping of customer keys, it would raise a novel regulatory question. Do both the consumer-facing bitcoin business, as well as the service provider it uses to secure its data, need to be licensed? Double-licensing would substantially erode any cost-savings thanks to firm specialization, and would likely discourage a competitive market for business-to-business virtual currency security. The result would be higher fees for consumers as well as less security.

As a result, only one party should be licensed in such a situation: the consumer-facing business. The consumer-facing business holds itself out as a trusted intermediary to its customers who may not have the time, expertise, or caution necessary to effectively comparison shop or hedge against risks. A business-to-business Bitcoin firm, on the other hand, offers its security services to savvy institutions who have both the motivation and the capacity to aggressively comparison shop. In short, while market failures may prevent competition from effectively protecting individual consumers, a competitive market unfettered by regulatory costs in the business-to-business arena would best enhance security. Moreover, as long as the consumer-facing business is a regulated entity, the protections of a Money Transmitter or Virtual Currency Business license will remain in effect for consumers.

[46] *See infra*.

[47] Cold storage involves placing the majority of an institution's private keys in offline media, either disconnected computer memory like a thumb-drive, paper, or as memorized passphrases—a so-called brain bank. If keys are not stored on Internet-connected servers, then they can only be accessed by compromising either the individual with access to the key or the physical security surrounding the key. The attack surface could thus be minimized by limiting the number of employees with knowledge of or access to offline key storage, and storing the offline drives or slips of paper in safe-deposit boxes or guarded premises.

[48] Amazon, *AWS Case Study: Netflix*, http://aws.amazon.com/solutions/case-studies/netflix/ .



Such a carve-out has been the longstanding norm for companies that are the legal agent of licensed money transmitters.[49] Similarly, the Financial Crimes Enforcement Network (“FinCEN”) exempts merchant processors and banking intermediaries from duties under the Bank Secrecy Act because these entities are merely intermediaries between banks, which are heavily regulated entities.[50] FinCEN also exempts those who only provide “the delivery, communication, or network access services used by a money transmitter to support money transmission services.”[51] Virtual Currency regulations should include a similar exemption in order to promote the development of enhanced security tools and services.

I. “Non-Monetary Purposes” and “Nominal Amount”

The technology underlying decentralized virtual currencies has promising applications apart from the provision of money transmission services. Distributed ledgers (or “blockchains”) are used within virtual currencies in order to keep a shared, write-only, public record of *who* has been sent *how many* units. Such a ledger may also find use in any area where records need to be authoritative, irreversible, and public.

Several non-monetary blockchain projects are already underway. They include distributed systems for Internet domain name registration, identity and authorization services (*e.g.* Blockstack), and notary services (*e.g.* Proof of Existence). Other companies are finding ways to simplify the process of setting up a blockchain for uses specific to a particular client. Much as RedHat helps IBM develop web servers using a particular version of the open-source Linux operating system, a blockchain specialist (*e.g.* Eris LTD) might help an accounting firm develop a specialized accounting system using blockchains.

Although these uses may have nothing to do with the provision of a money transmission service to consumers, they may nonetheless employ microtransactions in order to time-stamp some form of tokenized data. For example, a tiny fraction of a bitcoin (worth far less than one cent) may be sent on behalf of a customer in order to irreversibly note the identity of that customer on a public blockchain. The transaction is not intended to be a means of sending or receiving value; it is merely a representation of information that would be difficult to spoof, a verifiable token.

States may fear that such an exemption would create a dangerous loophole: a business could effectively operate as a money transmitting intermediary without licensure as long as it claims that the transactions are merely representing non-monetary data. As long as these placeholder transactions are small in value, however, there would be no viable way to use

[49] *See* New York Banking Law § 641 (“[N]or shall any person engage in such business as an agent, except as an agent of a licensee.”).

[50] 31 C.F.R. § 1010.100(ff)(5)(ii) (“The term “money transmitter” shall not include a person that only: . . . (B) Acts as a payment processor to facilitate the purchase of, or payment of a bill for, a good or service through a clearance and settlement system by agreement with the creditor or seller; (C) Operates a clearance and settlement system or otherwise acts as an intermediary solely between BSA regulated institutions.”).

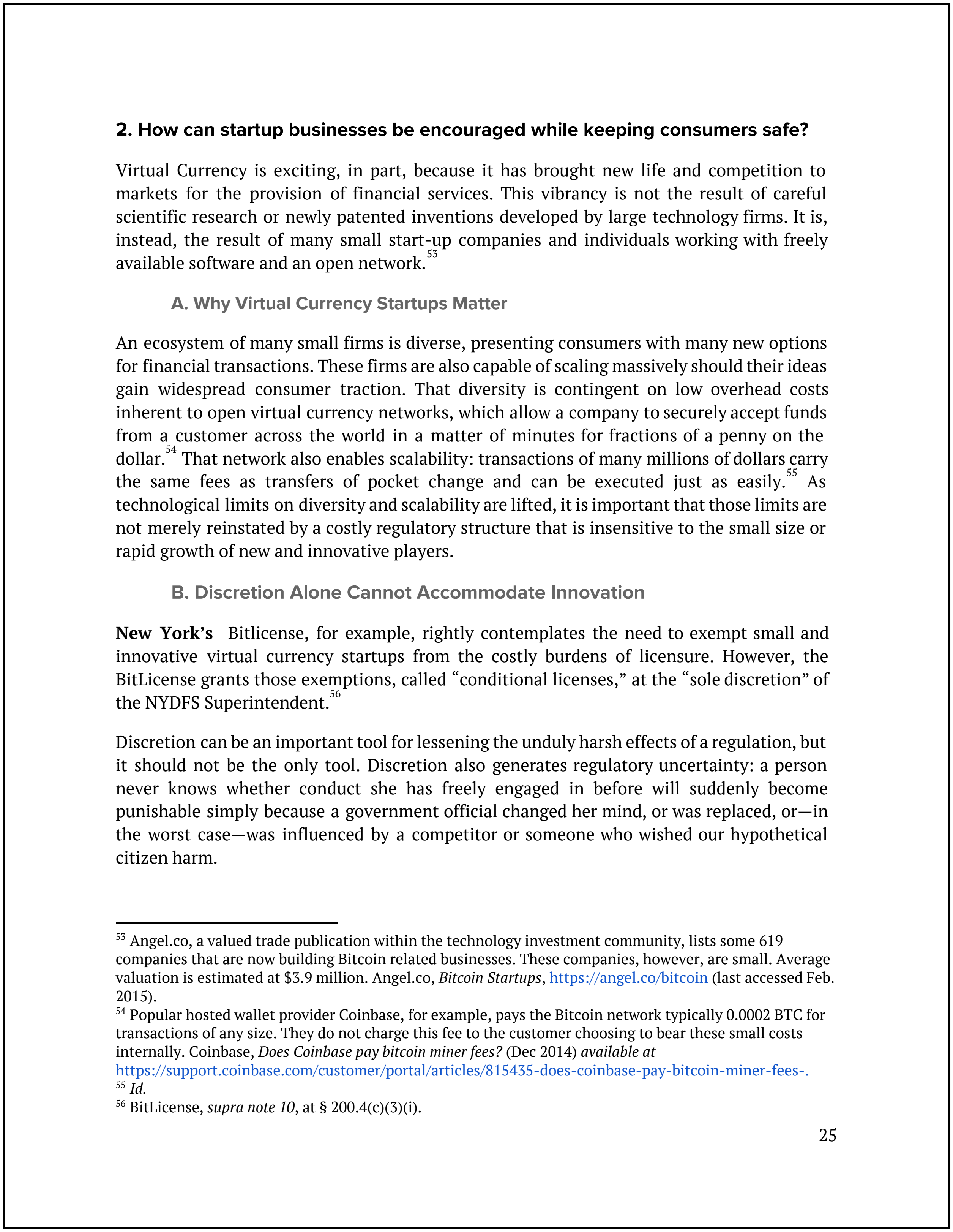
[51] 31 C.F.R. § 1010.100(ff)(5)(ii)(A).

such tools to transmit a meaningful amount of funds. As Muneeb Ali and Ryan Shea of Onename.io have explained:

> To illustrate with an example, if someone planned on moving $100 by breaking it up into 2,500 $0.04 transactions, they would have to pay a fee on the order of $0.04 for each and every transaction. Since moving the $100 from location A to location B would require 2,500 transactions to split up the money and 2,500 transactions to rejoin the money, the mover would be left with scattered denominations totaling $50 in the middle of the process and absolutely nothing by the end of the process. Second, if the mover ever wanted to reclaim all of those funds and make any use of them, they'd leave an enormous footprint on the blockchain, with thousands of suspicious addresses and transactions that people would be able to inspect and track. Thus, such transactions should be considered impractical for the movement of any kind of funds. It should be noted that any microtransaction that moves funds that are equal to or less than the minimum accepted network fee (today about $0.04), cannot possibly result in the transmission of any money whatsoever, as demonstrated above. Rather, they would result in the loss of 100% of funds by the time they are rejoined at the end of the process. By extension, orchestrated microtransactions that move funds equal to double the minimum accepted transaction fee would result in the loss of 50% of the total funds by the end of the process, and would still leave an enormous, conspicuous footprint.[52]

Accordingly, non-monetary transactions of *nominal* amounts should be outside the scope of Third Party Control of Virtual Currency regulation.

[52] Muneeb Ali & Ryan Shea, Comments to the New York Department of Financial Services on the Proposed Virtual Currency Regulatory Framework, *available at* http://www.dfs.ny.gov/legal/vcrf_0500/20141022%20VC%20Proposed%20Reg%20Comment%20245%20-%20OneName.pdf



2. How can startup businesses be encouraged while keeping consumers safe?

Virtual Currency is exciting, in part, because it has brought new life and competition to markets for the provision of financial services. This vibrancy is not the result of careful scientific research or newly patented inventions developed by large technology firms. It is, instead, the result of many small start-up companies and individuals working with freely available software and an open network.[53]

A. Why Virtual Currency Startups Matter

An ecosystem of many small firms is diverse, presenting consumers with many new options for financial transactions. These firms are also capable of scaling massively should their ideas gain widespread consumer traction. That diversity is contingent on low overhead costs inherent to open virtual currency networks, which allow a company to securely accept funds from a customer across the world in a matter of minutes for fractions of a penny on the dollar.[54] That network also enables scalability: transactions of many millions of dollars carry the same fees as transfers of pocket change and can be executed just as easily.[55] As technological limits on diversity and scalability are lifted, it is important that those limits are not merely reinstated by a costly regulatory structure that is insensitive to the small size or rapid growth of new and innovative players.

B. Discretion Alone Cannot Accommodate Innovation

New York's Bitlicense, for example, rightly contemplates the need to exempt small and innovative virtual currency startups from the costly burdens of licensure. However, the BitLicense grants those exemptions, called "conditional licenses," at the "sole discretion" of the NYDFS Superintendent.[56]

Discretion can be an important tool for lessening the unduly harsh effects of a regulation, but it should not be the only tool. Discretion also generates regulatory uncertainty: a person never knows whether conduct she has freely engaged in before will suddenly become punishable simply because a government official changed her mind, or was replaced, or—in the worst case—was influenced by a competitor or someone who wished our hypothetical citizen harm.

[53] Angel.co, a valued trade publication within the technology investment community, lists some 619 companies that are now building Bitcoin related businesses. These companies, however, are small. Average valuation is estimated at $3.9 million. Angel.co, *Bitcoin Startups*, https://angel.co/bitcoin (last accessed Feb. 2015).

[54] Popular hosted wallet provider Coinbase, for example, pays the Bitcoin network typically 0.0002 BTC for transactions of any size. They do not charge this fee to the customer choosing to bear these small costs internally. Coinbase, *Does Coinbase pay bitcoin miner fees?* (Dec 2014) *available at* https://support.coinbase.com/customer/portal/articles/815435-does-coinbase-pay-bitcoin-miner-fees-.

[55] *Id.*

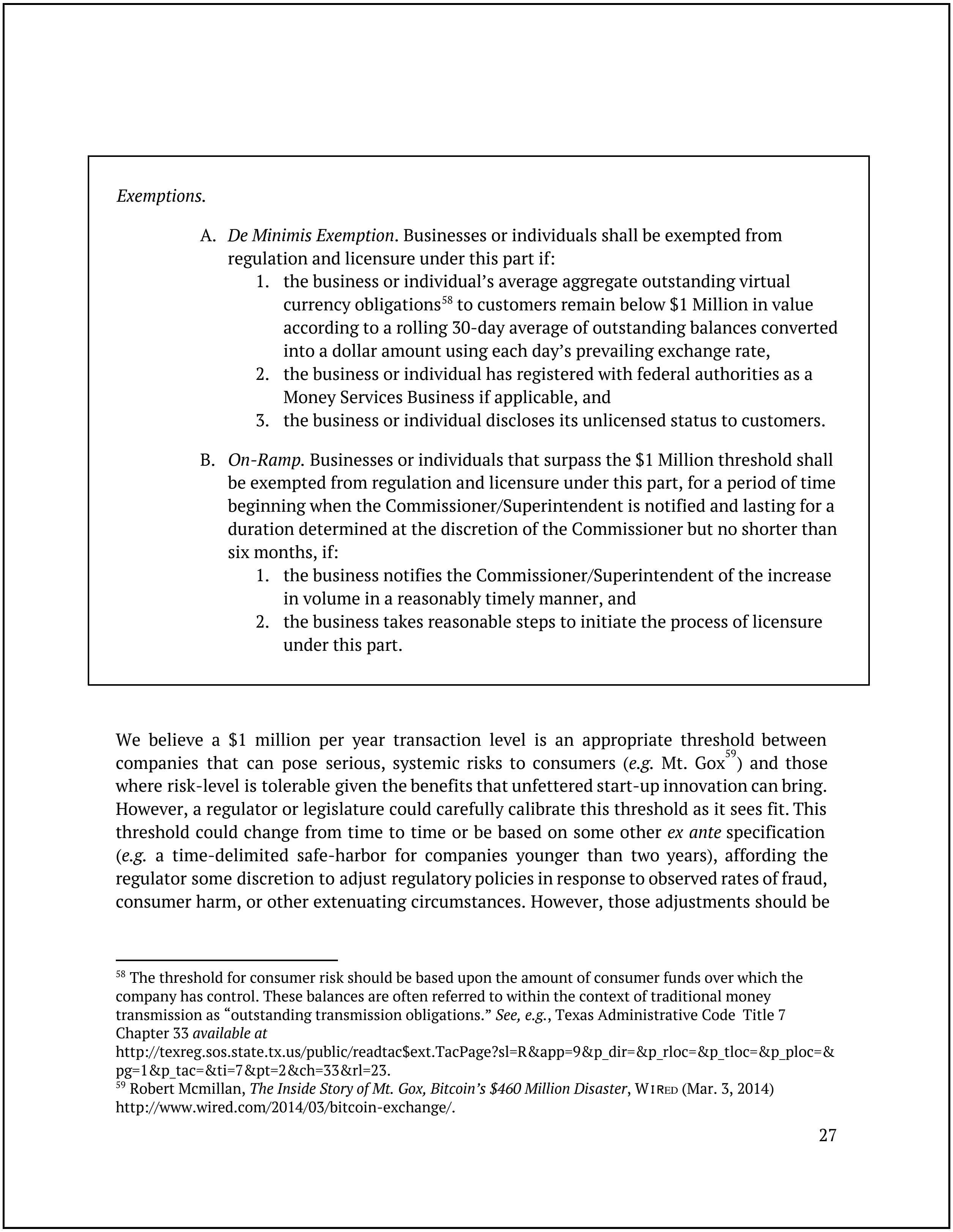
[56] BitLicense, *supra note 10*, at § 200.4(c)(3)(i).

A formal, rather than discretionary, carve-out for small startups is essential to preserve the freedom to innovate using these technologies, and it should be accomplished in a way that sets clear ex-ante standards and safe-harbors for budding entrepreneurs.

C. Drafting a De Minimis Exemption and On-Ramp for Startups

Small startups, academics, and hobbyists can be shielded from the costs of regulation and the severe criminal penalties that failure to license can trigger[57] by explicitly exempting them from regulation up until the point at which they pose non-trivial consumer risks; we can refer to this as a *de minimis exemption*. Shelter should also be granted to businesses that have passed that point and taken appropriate steps to alert the regulator and initiate the process of licensure; we can call this an *on-ramp* for startups. Language on the following page illustrates such exemptions.

[57] Courts are increasingly coming to the conclusion that virtual currencies such as Bitcoin qualify as "money" under various statutory definitions. Relatedly, any individual who "knowingly conducts, controls, manages, supervises, directs, or owns all or part" of a money services business operating without a money transmission license can be fined and imprisoned for up to five years under federal law. 18 U.S.C. §1960(a).



Exemptions.

A. *De Minimis Exemption*. Businesses or individuals shall be exempted from regulation and licensure under this part if:
   1. the business or individual's average aggregate outstanding virtual currency obligations[58] to customers remain below $1 Million in value according to a rolling 30-day average of outstanding balances converted into a dollar amount using each day's prevailing exchange rate,
   2. the business or individual has registered with federal authorities as a Money Services Business if applicable, and
   3. the business or individual discloses its unlicensed status to customers.

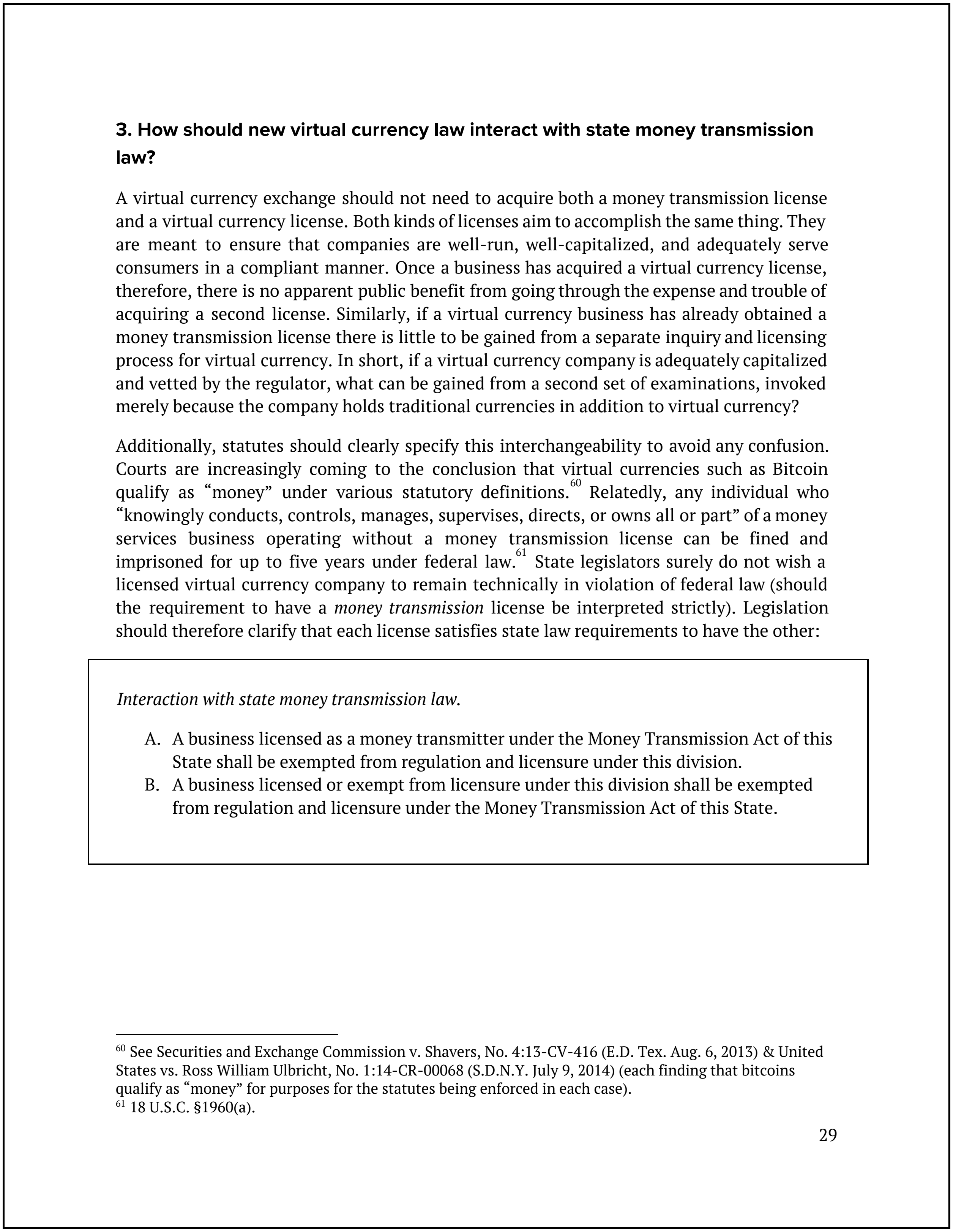
B. *On-Ramp.* Businesses or individuals that surpass the $1 Million threshold shall be exempted from regulation and licensure under this part, for a period of time beginning when the Commissioner/Superintendent is notified and lasting for a duration determined at the discretion of the Commissioner but no shorter than six months, if:
   1. the business notifies the Commissioner/Superintendent of the increase in volume in a reasonably timely manner, and
   2. the business takes reasonable steps to initiate the process of licensure under this part.

We believe a $1 million per year transaction level is an appropriate threshold between companies that can pose serious, systemic risks to consumers (*e.g.* Mt. Gox[59]) and those where risk-level is tolerable given the benefits that unfettered start-up innovation can bring. However, a regulator or legislature could carefully calibrate this threshold as it sees fit. This threshold could change from time to time or be based on some other *ex ante* specification (*e.g.* a time-delimited safe-harbor for companies younger than two years), affording the regulator some discretion to adjust regulatory policies in response to observed rates of fraud, consumer harm, or other extenuating circumstances. However, those adjustments should be

[58] The threshold for consumer risk should be based upon the amount of consumer funds over which the company has control. These balances are often referred to within the context of traditional money transmission as "outstanding transmission obligations." *See, e.g.*, Texas Administrative Code Title 7 Chapter 33 *available at* http://texreg.sos.state.tx.us/public/readtac$ext.TacPage?sl=R&app=9&p_dir=&p_rloc=&p_tloc=&p_ploc=&pg=1&p_tac=&ti=7&pt=2&ch=33&rl=23.

[59] Robert Mcmillan, *The Inside Story of Mt. Gox, Bitcoin's $460 Million Disaster*, Wired (Mar. 3, 2014) http://www.wired.com/2014/03/bitcoin-exchange/.

explicit, apply generally across the industry, and be announced in advance so that firms can plan their compliance strategies efficiently.



3. How should new virtual currency law interact with state money transmission law?

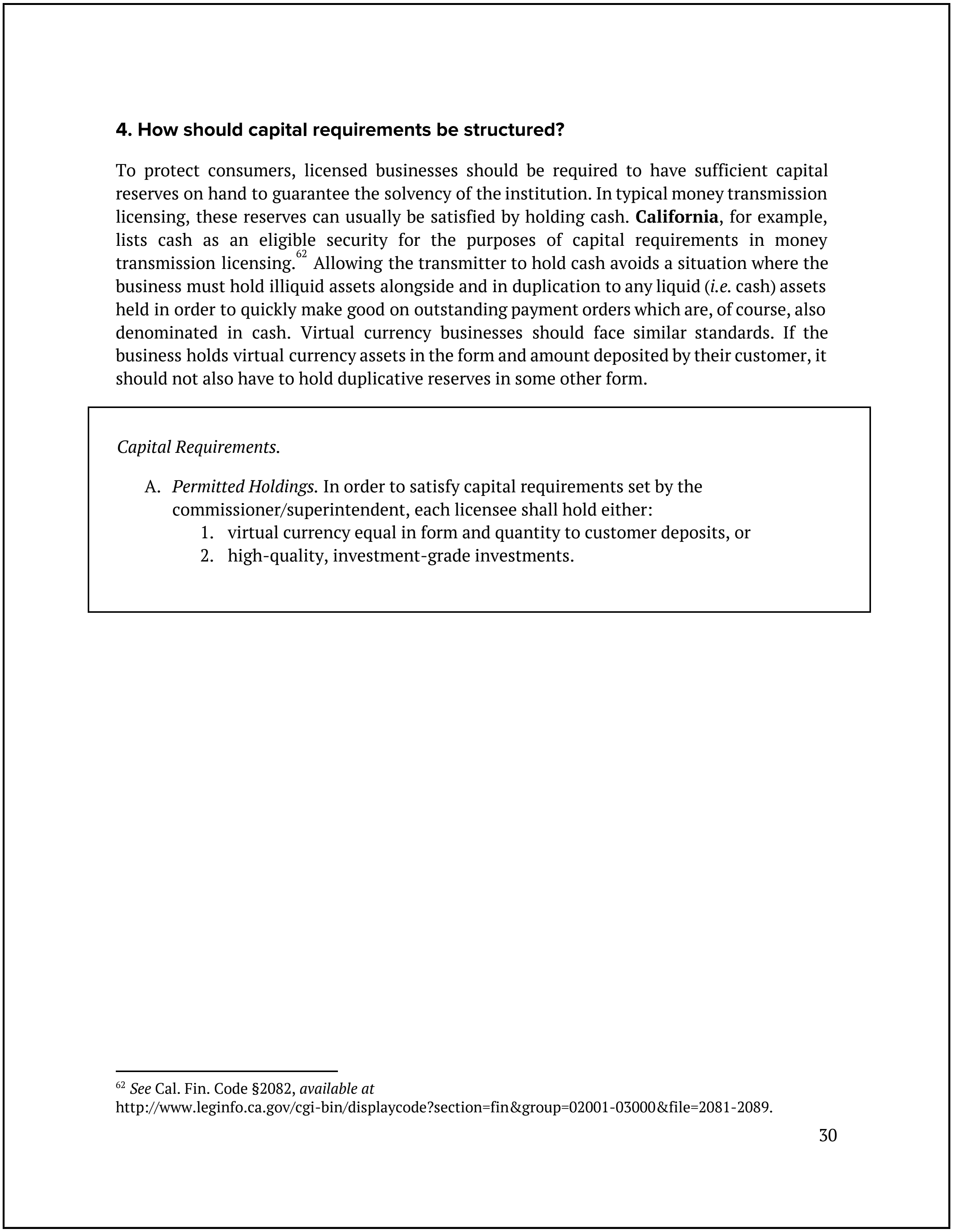
A virtual currency exchange should not need to acquire both a money transmission license and a virtual currency license. Both kinds of licenses aim to accomplish the same thing. They are meant to ensure that companies are well-run, well-capitalized, and adequately serve consumers in a compliant manner. Once a business has acquired a virtual currency license, therefore, there is no apparent public benefit from going through the expense and trouble of acquiring a second license. Similarly, if a virtual currency business has already obtained a money transmission license there is little to be gained from a separate inquiry and licensing process for virtual currency. In short, if a virtual currency company is adequately capitalized and vetted by the regulator, what can be gained from a second set of examinations, invoked merely because the company holds traditional currencies in addition to virtual currency?

Additionally, statutes should clearly specify this interchangeability to avoid any confusion. Courts are increasingly coming to the conclusion that virtual currencies such as Bitcoin qualify as "money" under various statutory definitions.[60] Relatedly, any individual who "knowingly conducts, controls, manages, supervises, directs, or owns all or part" of a money services business operating without a money transmission license can be fined and imprisoned for up to five years under federal law.[61] State legislators surely do not wish a licensed virtual currency company to remain technically in violation of federal law (should the requirement to have a *money transmission* license be interpreted strictly). Legislation should therefore clarify that each license satisfies state law requirements to have the other:

Interaction with state money transmission law.

A. A business licensed as a money transmitter under the Money Transmission Act of this State shall be exempted from regulation and licensure under this division.
B. A business licensed or exempt from licensure under this division shall be exempted from regulation and licensure under the Money Transmission Act of this State.

[60] See Securities and Exchange Commission v. Shavers, No. 4:13-CV-416 (E.D. Tex. Aug. 6, 2013) & United States vs. Ross William Ulbricht, No. 1:14-CR-00068 (S.D.N.Y. July 9, 2014) (each finding that bitcoins qualify as "money" for purposes for the statutes being enforced in each case).
[61] 18 U.S.C. §1960(a).

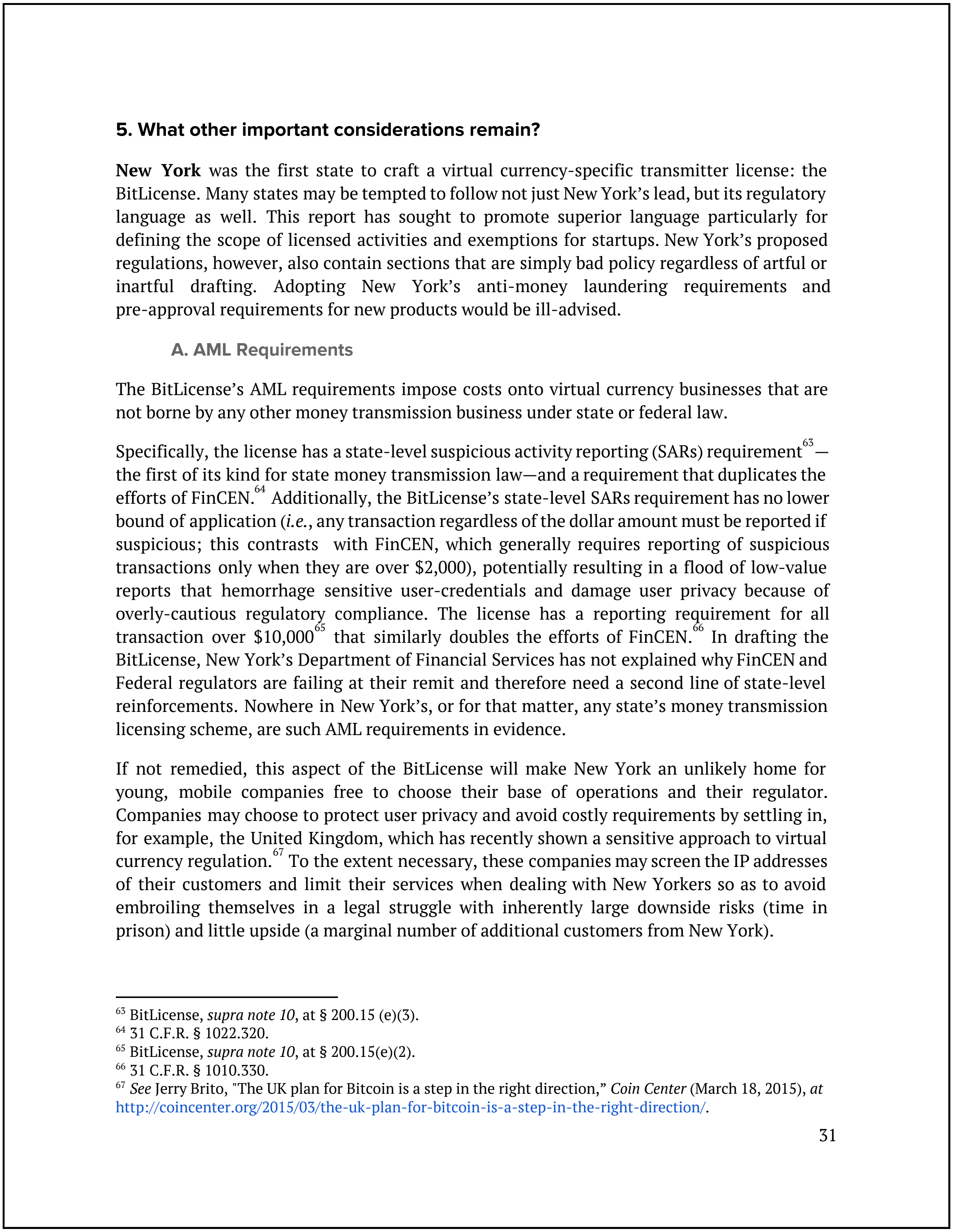


4. How should capital requirements be structured?

To protect consumers, licensed businesses should be required to have sufficient capital reserves on hand to guarantee the solvency of the institution. In typical money transmission licensing, these reserves can usually be satisfied by holding cash. **California**, for example, lists cash as an eligible security for the purposes of capital requirements in money transmission licensing.[62] Allowing the transmitter to hold cash avoids a situation where the business must hold illiquid assets alongside and in duplication to any liquid (*i.e.* cash) assets held in order to quickly make good on outstanding payment orders which are, of course, also denominated in cash. Virtual currency businesses should face similar standards. If the business holds virtual currency assets in the form and amount deposited by their customer, it should not also have to hold duplicative reserves in some other form.

> *Capital Requirements.*
>
> A. *Permitted Holdings.* In order to satisfy capital requirements set by the commissioner/superintendent, each licensee shall hold either:
>   1. virtual currency equal in form and quantity to customer deposits, or
>   2. high-quality, investment-grade investments.

[62] *See* Cal. Fin. Code §2082, *available at* http://www.leginfo.ca.gov/cgi-bin/displaycode?section=fin&group=02001-03000&file=2081-2089.



5. What other important considerations remain?

New York was the first state to craft a virtual currency-specific transmitter license: the BitLicense. Many states may be tempted to follow not just New York's lead, but its regulatory language as well. This report has sought to promote superior language particularly for defining the scope of licensed activities and exemptions for startups. New York's proposed regulations, however, also contain sections that are simply bad policy regardless of artful or inartful drafting. Adopting New York's anti-money laundering requirements and pre-approval requirements for new products would be ill-advised.

A. AML Requirements

The BitLicense's AML requirements impose costs onto virtual currency businesses that are not borne by any other money transmission business under state or federal law.

Specifically, the license has a state-level suspicious activity reporting (SARs) requirement[63]—the first of its kind for state money transmission law—and a requirement that duplicates the efforts of FinCEN.[64] Additionally, the BitLicense's state-level SARs requirement has no lower bound of application (*i.e.*, any transaction regardless of the dollar amount must be reported if suspicious; this contrasts with FinCEN, which generally requires reporting of suspicious transactions only when they are over $2,000), potentially resulting in a flood of low-value reports that hemorrhage sensitive user-credentials and damage user privacy because of overly-cautious regulatory compliance. The license has a reporting requirement for all transaction over $10,000[65] that similarly doubles the efforts of FinCEN.[66] In drafting the BitLicense, New York's Department of Financial Services has not explained why FinCEN and Federal regulators are failing at their remit and therefore need a second line of state-level reinforcements. Nowhere in New York's, or for that matter, any state's money transmission licensing scheme, are such AML requirements in evidence.

If not remedied, this aspect of the BitLicense will make New York an unlikely home for young, mobile companies free to choose their base of operations and their regulator. Companies may choose to protect user privacy and avoid costly requirements by settling in, for example, the United Kingdom, which has recently shown a sensitive approach to virtual currency regulation.[67] To the extent necessary, these companies may screen the IP addresses of their customers and limit their services when dealing with New Yorkers so as to avoid embroiling themselves in a legal struggle with inherently large downside risks (time in prison) and little upside (a marginal number of additional customers from New York).

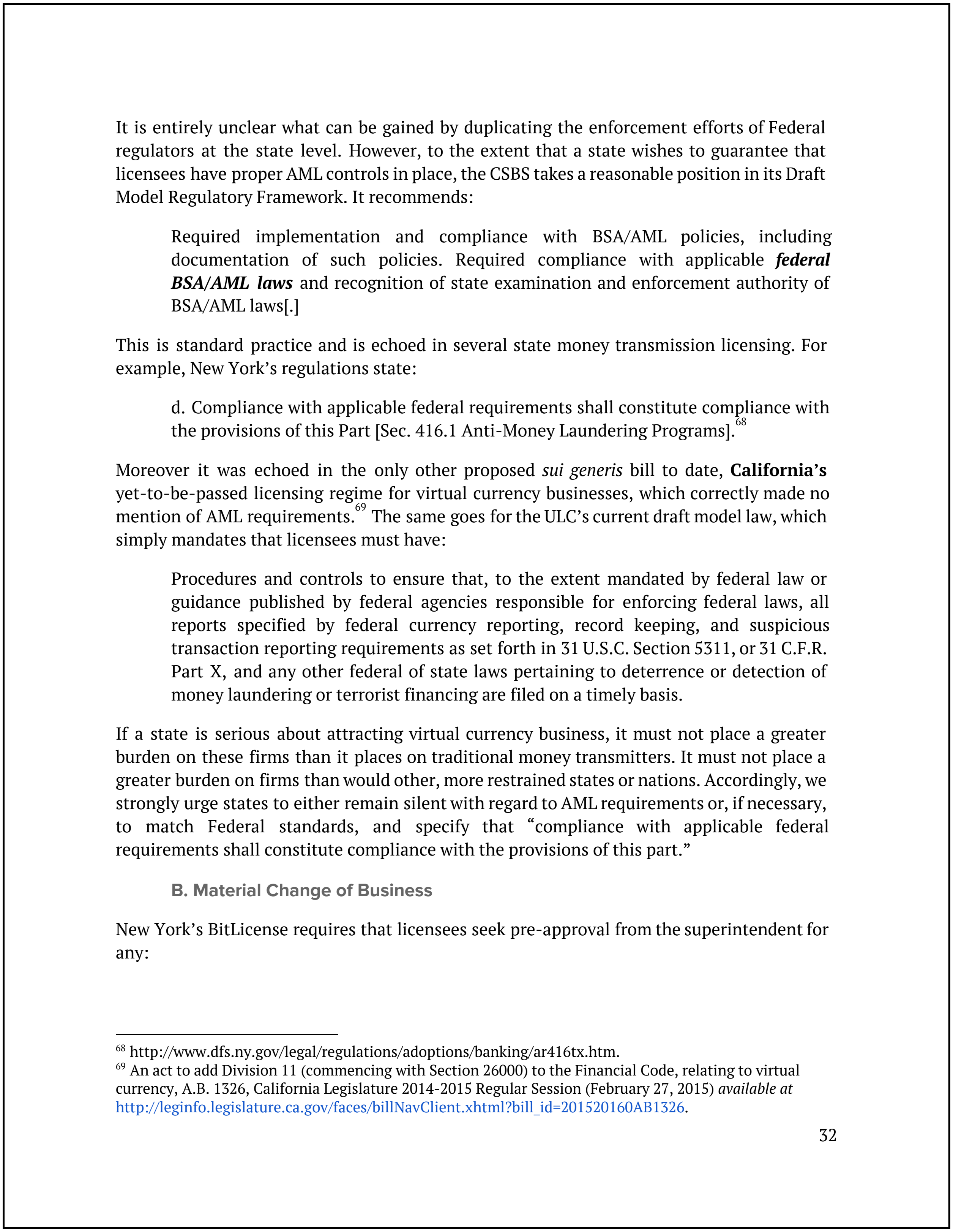
[63] BitLicense, *supra note 10*, at § 200.15 (e)(3).

[64] 31 C.F.R. § 1022.320.

[65] BitLicense, *supra note 10*, at § 200.15(e)(2).

[66] 31 C.F.R. § 1010.330.

[67] *See* Jerry Brito, "The UK plan for Bitcoin is a step in the right direction," *Coin Center* (March 18, 2015), *at* http://coincenter.org/2015/03/the-uk-plan-for-bitcoin-is-a-step-in-the-right-direction/.



It is entirely unclear what can be gained by duplicating the enforcement efforts of Federal regulators at the state level. However, to the extent that a state wishes to guarantee that licensees have proper AML controls in place, the CSBS takes a reasonable position in its Draft Model Regulatory Framework. It recommends:

> Required implementation and compliance with BSA/AML policies, including documentation of such policies. Required compliance with applicable ***federal BSA/AML laws*** and recognition of state examination and enforcement authority of BSA/AML laws[.]

This is standard practice and is echoed in several state money transmission licensing. For example, New York's regulations state:

> d. Compliance with applicable federal requirements shall constitute compliance with the provisions of this Part [Sec. 416.1 Anti-Money Laundering Programs].[68]

Moreover it was echoed in the only other proposed *sui generis* bill to date, **California's** yet-to-be-passed licensing regime for virtual currency businesses, which correctly made no mention of AML requirements.[69] The same goes for the ULC's current draft model law, which simply mandates that licensees must have:

> Procedures and controls to ensure that, to the extent mandated by federal law or guidance published by federal agencies responsible for enforcing federal laws, all reports specified by federal currency reporting, record keeping, and suspicious transaction reporting requirements as set forth in 31 U.S.C. Section 5311, or 31 C.F.R. Part X, and any other federal of state laws pertaining to deterrence or detection of money laundering or terrorist financing are filed on a timely basis.

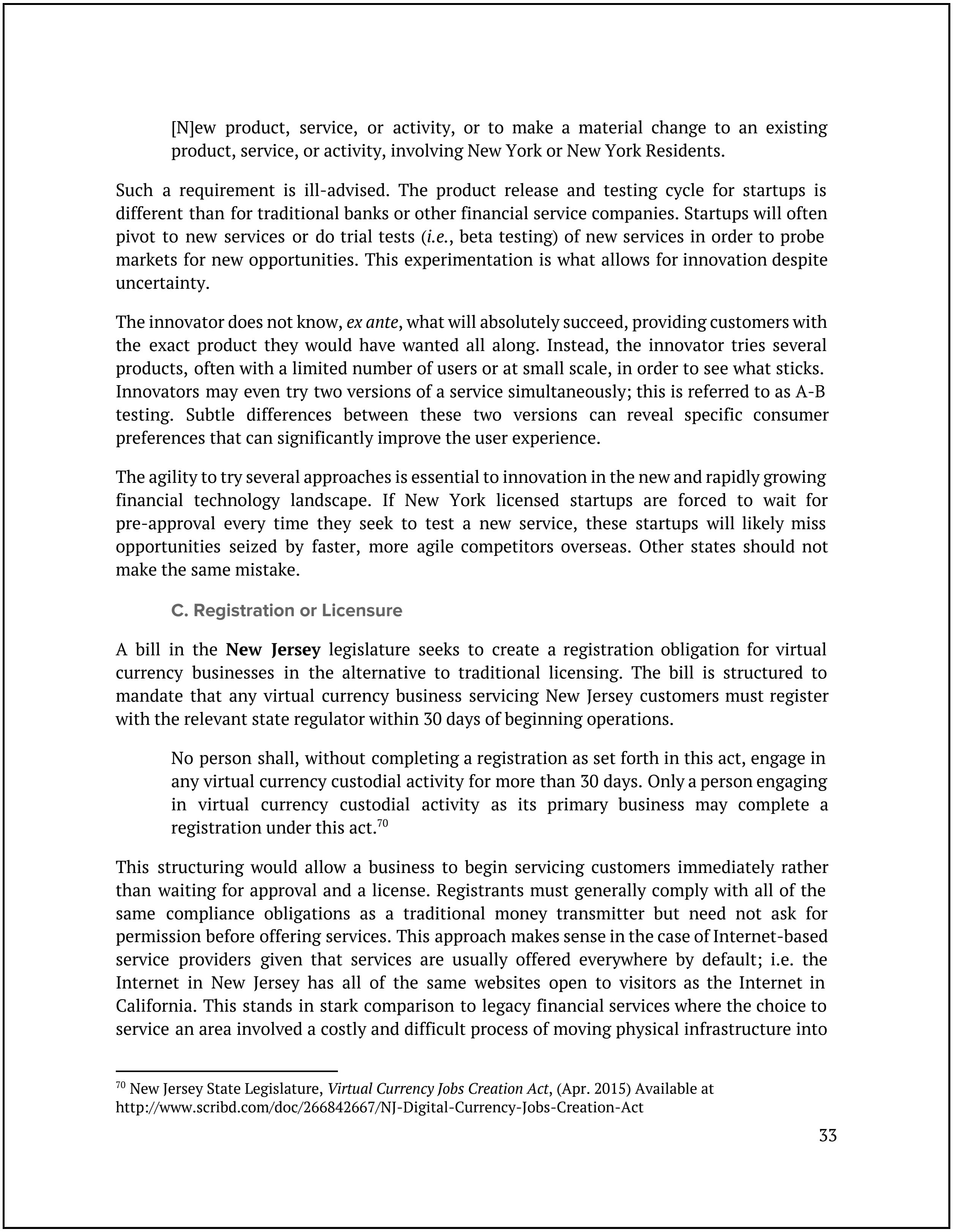
If a state is serious about attracting virtual currency business, it must not place a greater burden on these firms than it places on traditional money transmitters. It must not place a greater burden on firms than would other, more restrained states or nations. Accordingly, we strongly urge states to either remain silent with regard to AML requirements or, if necessary, to match Federal standards, and specify that "compliance with applicable federal requirements shall constitute compliance with the provisions of this part."

B. Material Change of Business

New York's BitLicense requires that licensees seek pre-approval from the superintendent for any:

[68] http://www.dfs.ny.gov/legal/regulations/adoptions/banking/ar416tx.htm.

[69] An act to add Division 11 (commencing with Section 26000) to the Financial Code, relating to virtual currency, A.B. 1326, California Legislature 2014-2015 Regular Session (February 27, 2015) *available at* http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160AB1326.



> [N]ew product, service, or activity, or to make a material change to an existing product, service, or activity, involving New York or New York Residents.

Such a requirement is ill-advised. The product release and testing cycle for startups is different than for traditional banks or other financial service companies. Startups will often pivot to new services or do trial tests (*i.e.*, beta testing) of new services in order to probe markets for new opportunities. This experimentation is what allows for innovation despite uncertainty.

The innovator does not know, *ex ante*, what will absolutely succeed, providing customers with the exact product they would have wanted all along. Instead, the innovator tries several products, often with a limited number of users or at small scale, in order to see what sticks. Innovators may even try two versions of a service simultaneously; this is referred to as A-B testing. Subtle differences between these two versions can reveal specific consumer preferences that can significantly improve the user experience.

The agility to try several approaches is essential to innovation in the new and rapidly growing financial technology landscape. If New York licensed startups are forced to wait for pre-approval every time they seek to test a new service, these startups will likely miss opportunities seized by faster, more agile competitors overseas. Other states should not make the same mistake.

C. Registration or Licensure

A bill in the **New Jersey** legislature seeks to create a registration obligation for virtual currency businesses in the alternative to traditional licensing. The bill is structured to mandate that any virtual currency business servicing New Jersey customers must register with the relevant state regulator within 30 days of beginning operations.

> No person shall, without completing a registration as set forth in this act, engage in any virtual currency custodial activity for more than 30 days. Only a person engaging in virtual currency custodial activity as its primary business may complete a registration under this act.[70]

This structuring would allow a business to begin servicing customers immediately rather than waiting for approval and a license. Registrants must generally comply with all of the same compliance obligations as a traditional money transmitter but need not ask for permission before offering services. This approach makes sense in the case of Internet-based service providers given that services are usually offered everywhere by default; i.e. the Internet in New Jersey has all of the same websites open to visitors as the Internet in California. This stands in stark comparison to legacy financial services where the choice to service an area involved a costly and difficult process of moving physical infrastructure into

[70] New Jersey State Legislature, *Virtual Currency Jobs Creation Act*, (Apr. 2015) Available at http://www.scribd.com/doc/266842667/NJ-Digital-Currency-Jobs-Creation-Act

the region or, at least, finding and negotiating with local agents. Limiting or blocking one's online service in states where licenses are pending is a difficult technological feat. States that wish to be leaders in the virtual currency and financial technology space should consider a registration-based approach to save service providers the difficulty of fragmenting the availability of their service and lagging against competitors while licenses are pending.

Leading states may also wish to consider offering tax-breaks to innovative companies, as are also proposed in the New Jersey bill.

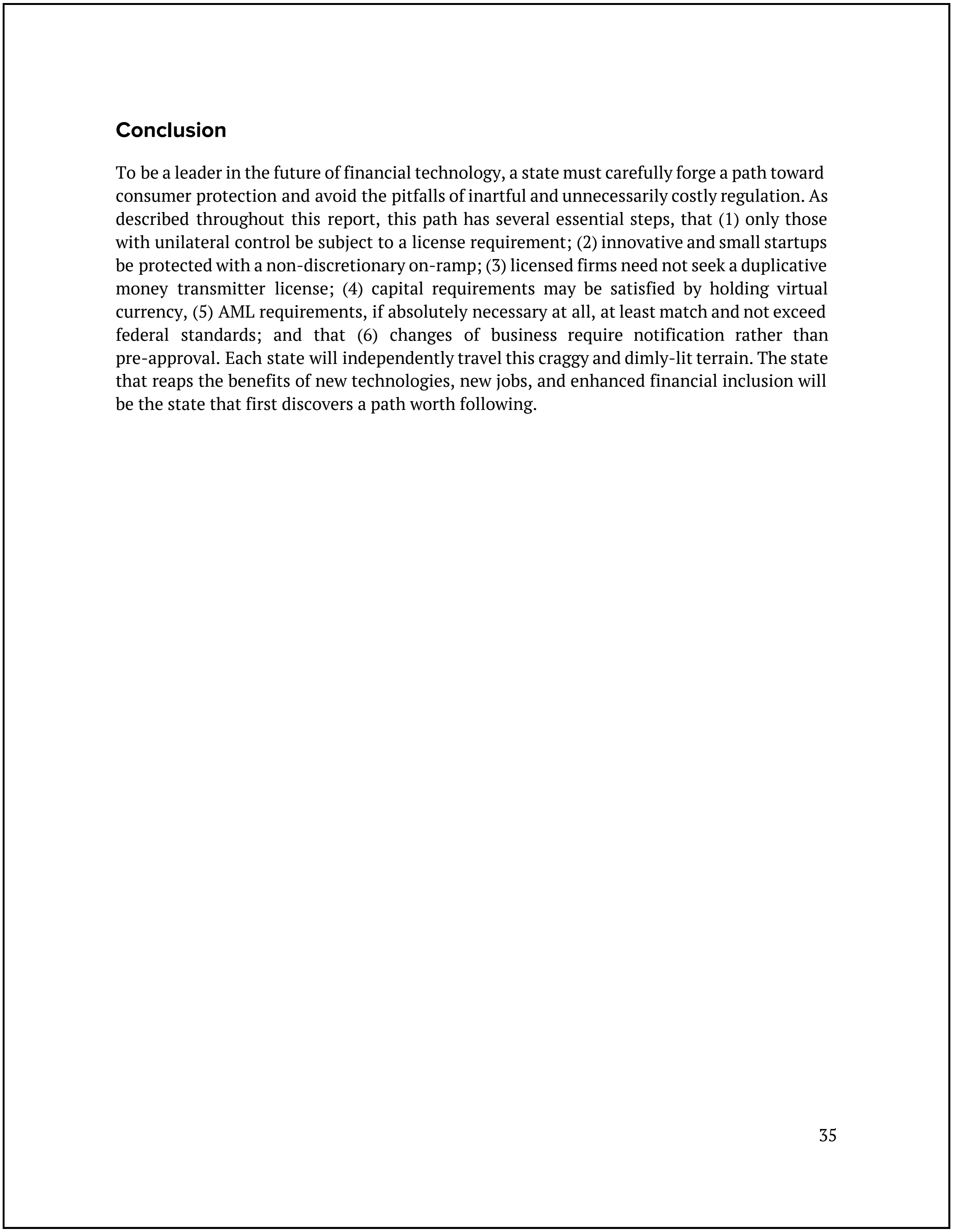
D. Agent of the Payee Exemption

Several states have formalized exemptions in money transmission law for so-called "agents of the payee."[71] At minimum, a state offering such an exemption to traditional money transmitters should treat virtual currency payment processors similarly. Additionally, there are some states where no formal exemption exists in the statute, but state regulators may consistently interpret their laws as not including agents of the payee. States taking this interpretive approach should consider crafting a formal exemption in the case of *sui generis* virtual currency legislation. Payee Agent Transactions should be exempted from licensing and defined as follows,

> Payee Agent Transactions. Transactions in which the recipient of virtual currency is an agent of the payee pursuant to a preexisting written contract and delivery of the virtual currency to the agent satisfies the payor's obligation to the payee.

or else the exemption should mirror existing language in the state's money transmission statute.

[71] California - SEC. 3. Section 2010 of the Financial Code: "This division does not apply to the following: ... (l) A transaction in which the recipient of the money or other monetary value is an agent of the payee pursuant to a preexisting written contract and delivery of the money or other monetary value to the agent satisfies the payor's obligation to the payee."
New York - Banking Law 641.1: "1. No person shall engage in the business of selling or issuing checks, or engage in the business of receiving money for transmission or transmitting the same, without a license therefor obtained from the superintendent as provided in this article, nor shall any person engage in such business as an agent, except as an agent of a licensee or as agent of a payee;"

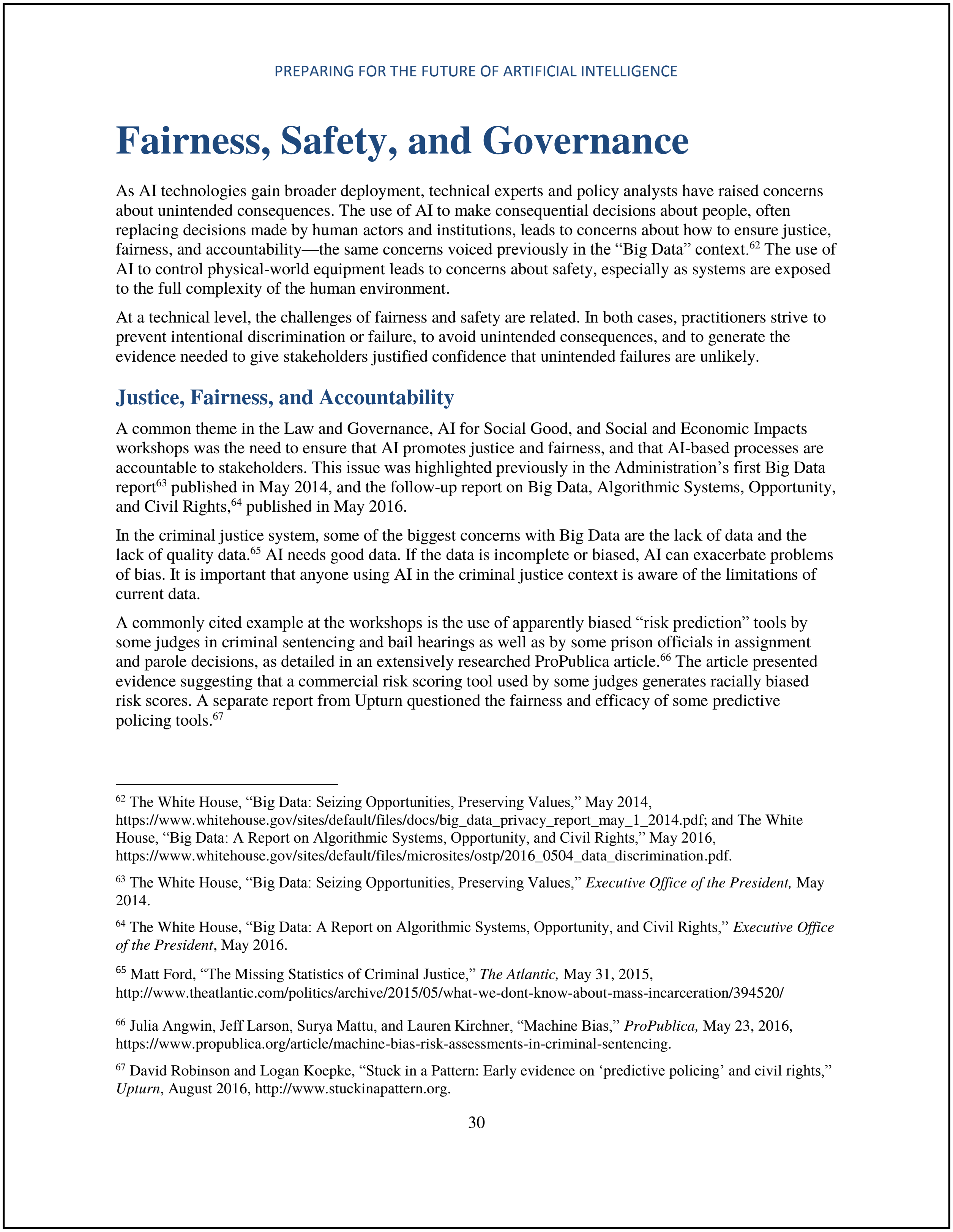


Conclusion

To be a leader in the future of financial technology, a state must carefully forge a path toward consumer protection and avoid the pitfalls of inartful and unnecessarily costly regulation. As described throughout this report, this path has several essential steps, that (1) only those with unilateral control be subject to a license requirement; (2) innovative and small startups be protected with a non-discretionary on-ramp; (3) licensed firms need not seek a duplicative money transmitter license; (4) capital requirements may be satisfied by holding virtual currency, (5) AML requirements, if absolutely necessary at all, at least match and not exceed federal standards; and that (6) changes of business require notification rather than pre-approval. Each state will independently travel this craggy and dimly-lit terrain. The state that reaps the benefits of new technologies, new jobs, and enhanced financial inclusion will be the state that first discovers a path worth following.

Chapter 21

Preparing for the Future of Artificial Intelligence, pp30-34 (NSTC)



Fairness, Safety, and Governance

As AI technologies gain broader deployment, technical experts and policy analysts have raised concerns about unintended consequences. The use of AI to make consequential decisions about people, often replacing decisions made by human actors and institutions, leads to concerns about how to ensure justice, fairness, and accountability—the same concerns voiced previously in the "Big Data" context.[62] The use of AI to control physical-world equipment leads to concerns about safety, especially as systems are exposed to the full complexity of the human environment.

At a technical level, the challenges of fairness and safety are related. In both cases, practitioners strive to prevent intentional discrimination or failure, to avoid unintended consequences, and to generate the evidence needed to give stakeholders justified confidence that unintended failures are unlikely.

Justice, Fairness, and Accountability

A common theme in the Law and Governance, AI for Social Good, and Social and Economic Impacts workshops was the need to ensure that AI promotes justice and fairness, and that AI-based processes are accountable to stakeholders. This issue was highlighted previously in the Administration's first Big Data report[63] published in May 2014, and the follow-up report on Big Data, Algorithmic Systems, Opportunity, and Civil Rights,[64] published in May 2016.

In the criminal justice system, some of the biggest concerns with Big Data are the lack of data and the lack of quality data.[65] AI needs good data. If the data is incomplete or biased, AI can exacerbate problems of bias. It is important that anyone using AI in the criminal justice context is aware of the limitations of current data.

A commonly cited example at the workshops is the use of apparently biased "risk prediction" tools by some judges in criminal sentencing and bail hearings as well as by some prison officials in assignment and parole decisions, as detailed in an extensively researched ProPublica article.[66] The article presented evidence suggesting that a commercial risk scoring tool used by some judges generates racially biased risk scores. A separate report from Upturn questioned the fairness and efficacy of some predictive policing tools.[67]

[62] The White House, "Big Data: Seizing Opportunities, Preserving Values," May 2014, https://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf; and The White House, "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights," May 2016, https://www.whitehouse.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf.

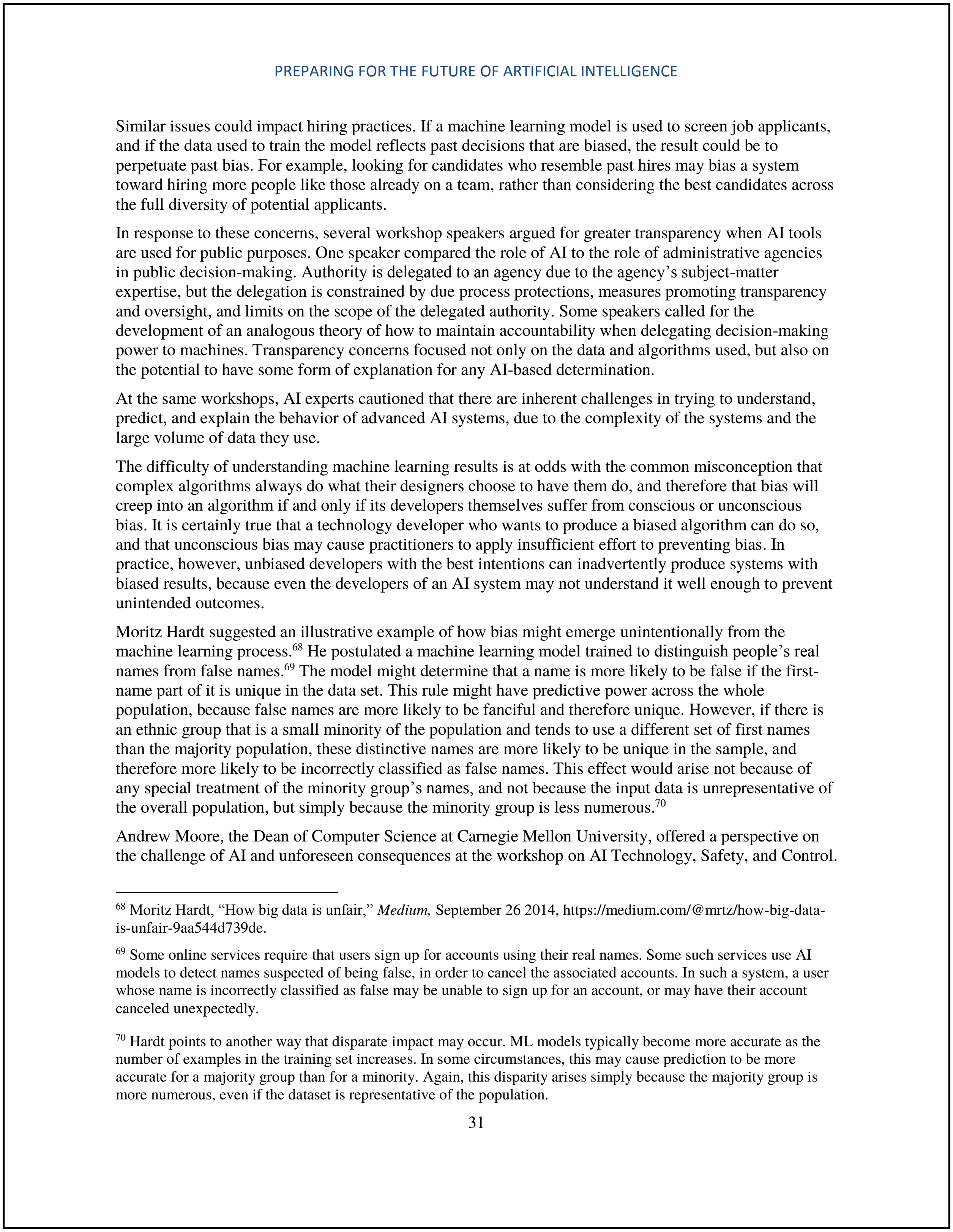
[63] The White House, "Big Data: Seizing Opportunities, Preserving Values," *Executive Office of the President,* May 2014.

[64] The White House, "Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights," *Executive Office of the President*, May 2016.

[65] Matt Ford, "The Missing Statistics of Criminal Justice," *The Atlantic,* May 31, 2015, http://www.theatlantic.com/politics/archive/2015/05/what-we-dont-know-about-mass-incarceration/394520/

[66] Julia Angwin, Jeff Larson, Surya Mattu, and Lauren Kirchner, "Machine Bias," *ProPublica,* May 23, 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.

[67] David Robinson and Logan Koepke, "Stuck in a Pattern: Early evidence on 'predictive policing' and civil rights," *Upturn*, August 2016, http://www.stuckinapattern.org.



Similar issues could impact hiring practices. If a machine learning model is used to screen job applicants, and if the data used to train the model reflects past decisions that are biased, the result could be to perpetuate past bias. For example, looking for candidates who resemble past hires may bias a system toward hiring more people like those already on a team, rather than considering the best candidates across the full diversity of potential applicants.

In response to these concerns, several workshop speakers argued for greater transparency when AI tools are used for public purposes. One speaker compared the role of AI to the role of administrative agencies in public decision-making. Authority is delegated to an agency due to the agency's subject-matter expertise, but the delegation is constrained by due process protections, measures promoting transparency and oversight, and limits on the scope of the delegated authority. Some speakers called for the development of an analogous theory of how to maintain accountability when delegating decision-making power to machines. Transparency concerns focused not only on the data and algorithms used, but also on the potential to have some form of explanation for any AI-based determination.

At the same workshops, AI experts cautioned that there are inherent challenges in trying to understand, predict, and explain the behavior of advanced AI systems, due to the complexity of the systems and the large volume of data they use.

The difficulty of understanding machine learning results is at odds with the common misconception that complex algorithms always do what their designers choose to have them do, and therefore that bias will creep into an algorithm if and only if its developers themselves suffer from conscious or unconscious bias. It is certainly true that a technology developer who wants to produce a biased algorithm can do so, and that unconscious bias may cause practitioners to apply insufficient effort to preventing bias. In practice, however, unbiased developers with the best intentions can inadvertently produce systems with biased results, because even the developers of an AI system may not understand it well enough to prevent unintended outcomes.

Moritz Hardt suggested an illustrative example of how bias might emerge unintentionally from the machine learning process.[68] He postulated a machine learning model trained to distinguish people's real names from false names.[69] The model might determine that a name is more likely to be false if the first-name part of it is unique in the data set. This rule might have predictive power across the whole population, because false names are more likely to be fanciful and therefore unique. However, if there is an ethnic group that is a small minority of the population and tends to use a different set of first names than the majority population, these distinctive names are more likely to be unique in the sample, and therefore more likely to be incorrectly classified as false names. This effect would arise not because of any special treatment of the minority group's names, and not because the input data is unrepresentative of the overall population, but simply because the minority group is less numerous.[70]

Andrew Moore, the Dean of Computer Science at Carnegie Mellon University, offered a perspective on the challenge of AI and unforeseen consequences at the workshop on AI Technology, Safety, and Control.

[68] Moritz Hardt, "How big data is unfair," *Medium,* September 26 2014, https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de.

[69] Some online services require that users sign up for accounts using their real names. Some such services use AI models to detect names suspected of being false, in order to cancel the associated accounts. In such a system, a user whose name is incorrectly classified as false may be unable to sign up for an account, or may have their account canceled unexpectedly.

[70] Hardt points to another way that disparate impact may occur. ML models typically become more accurate as the number of examples in the training set increases. In some circumstances, this may cause prediction to be more accurate for a majority group than for a minority. Again, this disparity arises simply because the majority group is more numerous, even if the dataset is representative of the population.

He argued that today, because of the opacity of AI algorithms, the most effective way to minimize the risk of unintended outcomes is through extensive testing—essentially to make a long list of the types of bad outcomes that could occur, and to rule out these outcomes by creating many specialized tests to look for them.

An example of what can go wrong in the absence of extensive testing comes from a trained model for automatically captioning photos, which infamously put the caption “gorilla” on some close-up photos of dark-skinned human faces. This was antithetical to the developers’ values, and it occurred despite testing that showed the model produced accurate results on a high percentage of all photos. These particular errors, although rare, had negative consequences that were beyond the understanding of the model, which had no built-in concept of race, nor any understanding of the relevant historical context. One way to prevent this type of error would have involved extensive testing of the algorithm to scrutinize how human faces, in particular, are labeled, including examination of some results by people who could recognize unacceptable outcomes that the model wouldn’t catch.

Ethical training for AI practitioners and students is a necessary part of the solution. Ideally, every student learning AI, computer science, or data science would be exposed to curriculum and discussion on related ethics and security topics.[71] However, ethics alone is not sufficient. Ethics can help practitioners understand their responsibilities to all stakeholders, but ethical training needs to be augmented with the technical capability to put good intentions into practice by taking technical precautions as a system is built and tested.

As practitioners strive to make AI systems more just, fair and accountable, there are often opportunities to make technology an aid to accountability rather than a barrier to it. Research to improve the interpretability of machine learning results is one example. Having an interpretable model that helps people understand a decision empowers them to interrogate the assumptions and processes behind it.

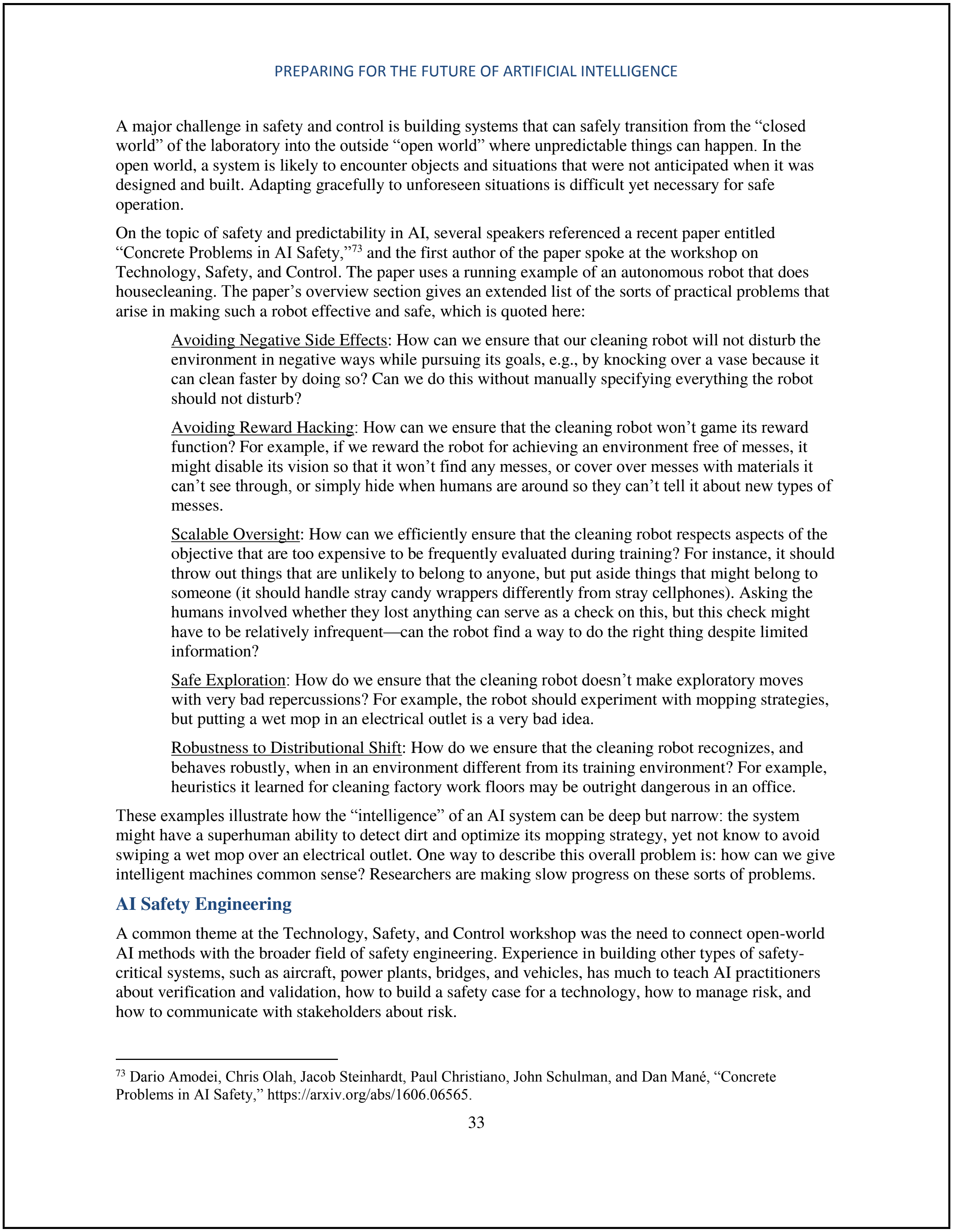
There are several technical approaches to enhancing the accountability and robustness of complex algorithmic decisions. A system can be tested “in the wild” by presenting it with situations and observing its behavior. A system can be subjected to black-box testing, in which it is presented with synthetic inputs and its behavior is observed, enabling behavior to be tested in scenarios that might not occur naturally.[72] Some or all of the technical details of a system’s design can be published, enabling analysts to replicate it and analyze aspects of its internal behavior that might be difficult to characterize with testing alone. In some cases it is possible to publish information that helps the public evaluate a system’s risk of bias, while withholding other information about the system as proprietary or private.

Safety and Control

At the workshops, AI experts said that one of the main factors limiting the deployment of AI in the real world is concern about safety and control. If practitioners cannot achieve justified confidence that a system is safe and controllable, so that deploying the system does not create an unacceptable risk of serious negative consequences, then the system cannot and should not be deployed.

[71] Some institutions may choose to incorporate ethics into existing courses. Others may choose to introduce separate courses on ethics.

[72] Black-box testing allows a system to be presented with fictionalized data, which enables comprehensive experiments that vary individual attributes of an individual as well as larger numbers of experiments than might be possible for in-the-wild testing. See, e.g., Anupam Datta, Shayak Sen, and Yair Zick, “Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems,” *Proceedings of 37th IEEE Symposium on Security and Privacy*, 2016.



A major challenge in safety and control is building systems that can safely transition from the "closed world" of the laboratory into the outside "open world" where unpredictable things can happen. In the open world, a system is likely to encounter objects and situations that were not anticipated when it was designed and built. Adapting gracefully to unforeseen situations is difficult yet necessary for safe operation.

On the topic of safety and predictability in AI, several speakers referenced a recent paper entitled "Concrete Problems in AI Safety,"[73] and the first author of the paper spoke at the workshop on Technology, Safety, and Control. The paper uses a running example of an autonomous robot that does housecleaning. The paper's overview section gives an extended list of the sorts of practical problems that arise in making such a robot effective and safe, which is quoted here:

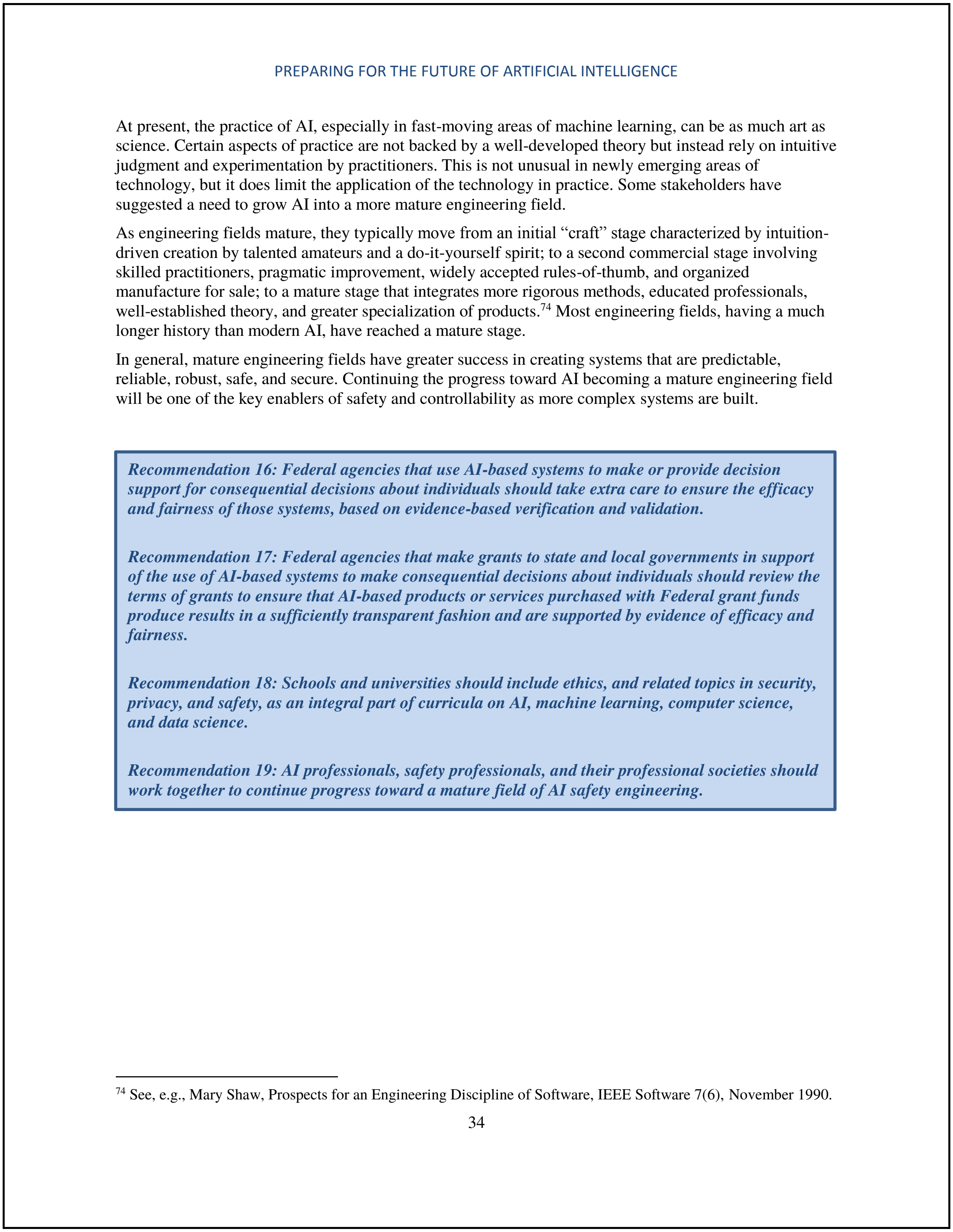
> Avoiding Negative Side Effects: How can we ensure that our cleaning robot will not disturb the environment in negative ways while pursuing its goals, e.g., by knocking over a vase because it can clean faster by doing so? Can we do this without manually specifying everything the robot should not disturb?
>
> Avoiding Reward Hacking: How can we ensure that the cleaning robot won't game its reward function? For example, if we reward the robot for achieving an environment free of messes, it might disable its vision so that it won't find any messes, or cover over messes with materials it can't see through, or simply hide when humans are around so they can't tell it about new types of messes.
>
> Scalable Oversight: How can we efficiently ensure that the cleaning robot respects aspects of the objective that are too expensive to be frequently evaluated during training? For instance, it should throw out things that are unlikely to belong to anyone, but put aside things that might belong to someone (it should handle stray candy wrappers differently from stray cellphones). Asking the humans involved whether they lost anything can serve as a check on this, but this check might have to be relatively infrequent—can the robot find a way to do the right thing despite limited information?
>
> Safe Exploration: How do we ensure that the cleaning robot doesn't make exploratory moves with very bad repercussions? For example, the robot should experiment with mopping strategies, but putting a wet mop in an electrical outlet is a very bad idea.
>
> Robustness to Distributional Shift: How do we ensure that the cleaning robot recognizes, and behaves robustly, when in an environment different from its training environment? For example, heuristics it learned for cleaning factory work floors may be outright dangerous in an office.

These examples illustrate how the "intelligence" of an AI system can be deep but narrow: the system might have a superhuman ability to detect dirt and optimize its mopping strategy, yet not know to avoid swiping a wet mop over an electrical outlet. One way to describe this overall problem is: how can we give intelligent machines common sense? Researchers are making slow progress on these sorts of problems.

AI Safety Engineering

A common theme at the Technology, Safety, and Control workshop was the need to connect open-world AI methods with the broader field of safety engineering. Experience in building other types of safety-critical systems, such as aircraft, power plants, bridges, and vehicles, has much to teach AI practitioners about verification and validation, how to build a safety case for a technology, how to manage risk, and how to communicate with stakeholders about risk.

[73] Dario Amodei, Chris Olah, Jacob Steinhardt, Paul Christiano, John Schulman, and Dan Mané, "Concrete Problems in AI Safety," https://arxiv.org/abs/1606.06565.



At present, the practice of AI, especially in fast-moving areas of machine learning, can be as much art as science. Certain aspects of practice are not backed by a well-developed theory but instead rely on intuitive judgment and experimentation by practitioners. This is not unusual in newly emerging areas of technology, but it does limit the application of the technology in practice. Some stakeholders have suggested a need to grow AI into a more mature engineering field.

As engineering fields mature, they typically move from an initial "craft" stage characterized by intuition-driven creation by talented amateurs and a do-it-yourself spirit; to a second commercial stage involving skilled practitioners, pragmatic improvement, widely accepted rules-of-thumb, and organized manufacture for sale; to a mature stage that integrates more rigorous methods, educated professionals, well-established theory, and greater specialization of products.[74] Most engineering fields, having a much longer history than modern AI, have reached a mature stage.

In general, mature engineering fields have greater success in creating systems that are predictable, reliable, robust, safe, and secure. Continuing the progress toward AI becoming a mature engineering field will be one of the key enablers of safety and controllability as more complex systems are built.

Recommendation 16: Federal agencies that use AI-based systems to make or provide decision support for consequential decisions about individuals should take extra care to ensure the efficacy and fairness of those systems, based on evidence-based verification and validation.

Recommendation 17: Federal agencies that make grants to state and local governments in support of the use of AI-based systems to make consequential decisions about individuals should review the terms of grants to ensure that AI-based products or services purchased with Federal grant funds produce results in a sufficiently transparent fashion and are supported by evidence of efficacy and fairness.

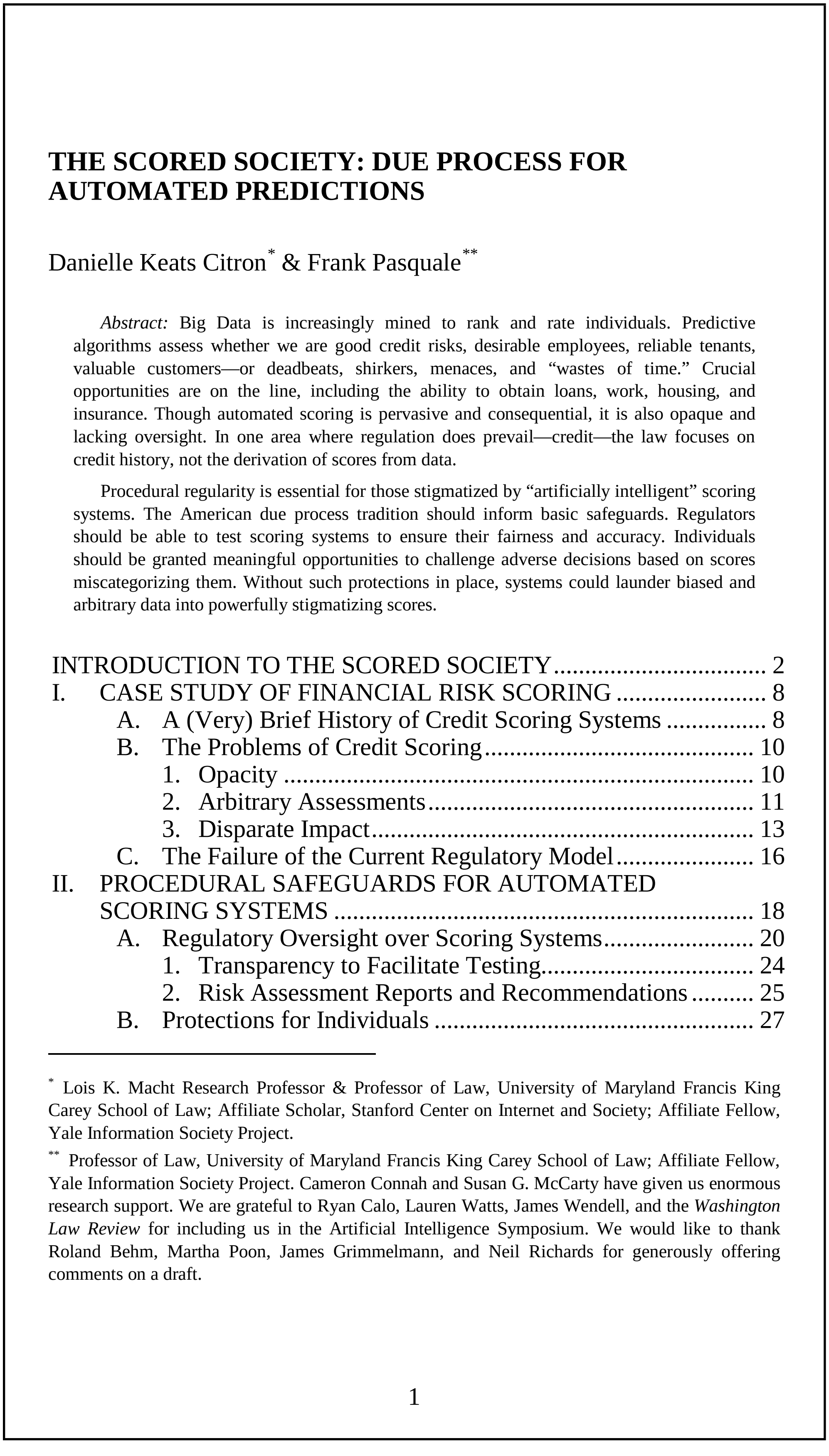
Recommendation 18: Schools and universities should include ethics, and related topics in security, privacy, and safety, as an integral part of curricula on AI, machine learning, computer science, and data science.

Recommendation 19: AI professionals, safety professionals, and their professional societies should work together to continue progress toward a mature field of AI safety engineering.

[74] See, e.g., Mary Shaw, Prospects for an Engineering Discipline of Software, IEEE Software 7(6), November 1990.

Chapter 22

The Scored Society: Due Process for Automated Predictions (Citron and Pasquale)



THE SCORED SOCIETY: DUE PROCESS FOR AUTOMATED PREDICTIONS

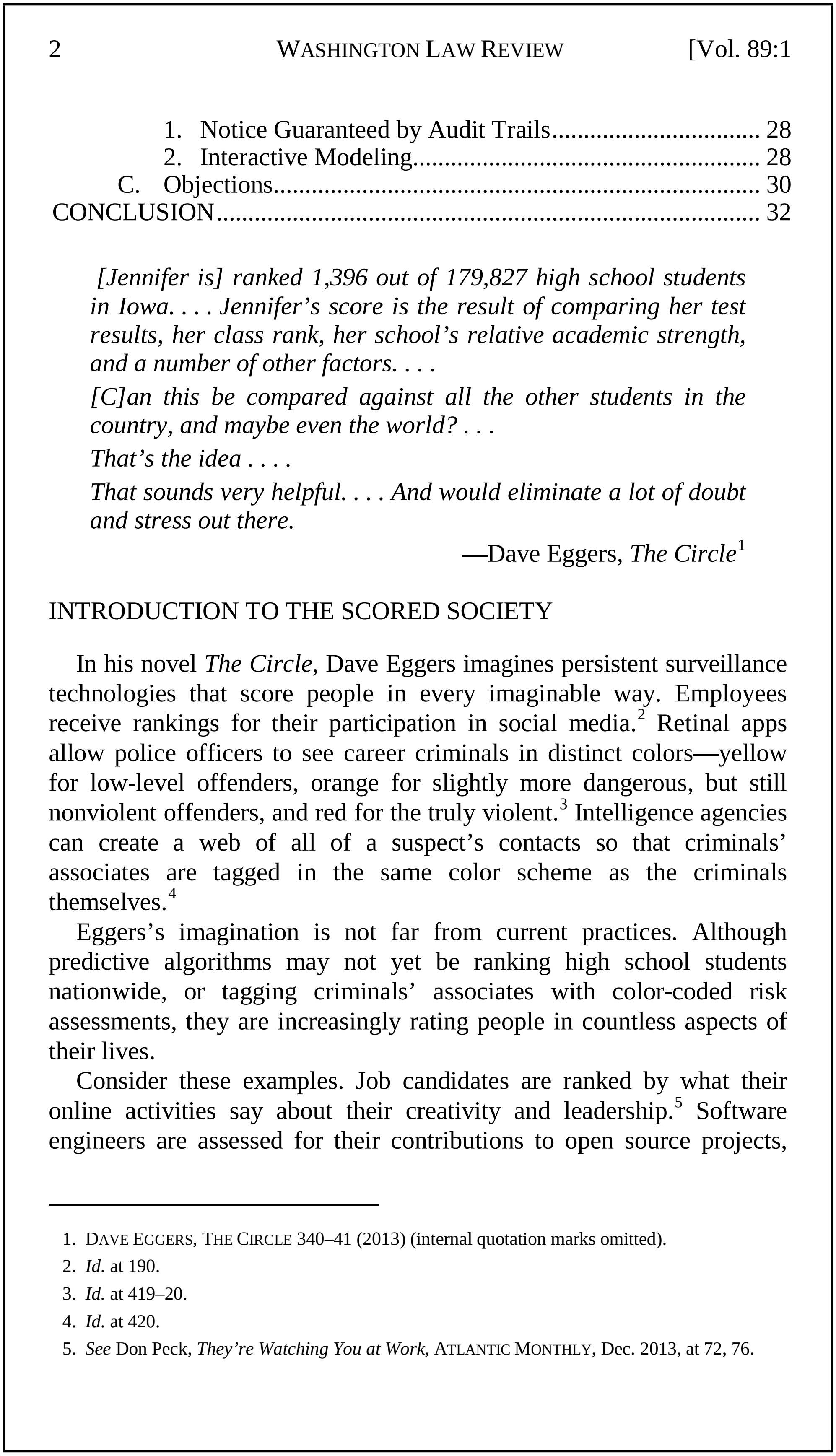
Danielle Keats Citron[*] & Frank Pasquale[**]

Abstract: Big Data is increasingly mined to rank and rate individuals. Predictive algorithms assess whether we are good credit risks, desirable employees, reliable tenants, valuable customers—or deadbeats, shirkers, menaces, and "wastes of time." Crucial opportunities are on the line, including the ability to obtain loans, work, housing, and insurance. Though automated scoring is pervasive and consequential, it is also opaque and lacking oversight. In one area where regulation does prevail—credit—the law focuses on credit history, not the derivation of scores from data.

Procedural regularity is essential for those stigmatized by "artificially intelligent" scoring systems. The American due process tradition should inform basic safeguards. Regulators should be able to test scoring systems to ensure their fairness and accuracy. Individuals should be granted meaningful opportunities to challenge adverse decisions based on scores miscategorizing them. Without such protections in place, systems could launder biased and arbitrary data into powerfully stigmatizing scores.

[*] Lois K. Macht Research Professor & Professor of Law, University of Maryland Francis King Carey School of Law; Affiliate Scholar, Stanford Center on Internet and Society; Affiliate Fellow, Yale Information Society Project.

[**] Professor of Law, University of Maryland Francis King Carey School of Law; Affiliate Fellow, Yale Information Society Project. Cameron Connah and Susan G. McCarty have given us enormous research support. We are grateful to Ryan Calo, Lauren Watts, James Wendell, and the *Washington Law Review* for including us in the Artificial Intelligence Symposium. We would like to thank Roland Behm, Martha Poon, James Grimmelmann, and Neil Richards for generously offering comments on a draft.



> *[Jennifer is] ranked 1,396 out of 179,827 high school students in Iowa. . . . Jennifer's score is the result of comparing her test results, her class rank, her school's relative academic strength, and a number of other factors. . . .*
>
> *[C]an this be compared against all the other students in the country, and maybe even the world? . . .*
>
> *That's the idea . . . .*
>
> *That sounds very helpful. . . . And would eliminate a lot of doubt and stress out there.*
>
> —Dave Eggers, *The Circle*[1]

INTRODUCTION TO THE SCORED SOCIETY

In his novel *The Circle*, Dave Eggers imagines persistent surveillance technologies that score people in every imaginable way. Employees receive rankings for their participation in social media.[2] Retinal apps allow police officers to see career criminals in distinct colors—yellow for low-level offenders, orange for slightly more dangerous, but still nonviolent offenders, and red for the truly violent.[3] Intelligence agencies can create a web of all of a suspect's contacts so that criminals' associates are tagged in the same color scheme as the criminals themselves.[4]

Eggers's imagination is not far from current practices. Although predictive algorithms may not yet be ranking high school students nationwide, or tagging criminals' associates with color-coded risk assessments, they are increasingly rating people in countless aspects of their lives.

Consider these examples. Job candidates are ranked by what their online activities say about their creativity and leadership.[5] Software engineers are assessed for their contributions to open source projects,

[1] DAVE EGGERS, THE CIRCLE 340–41 (2013) (internal quotation marks omitted).

[2] *Id.* at 190.

[3] *Id.* at 419–20.

[4] *Id.* at 420.

[5] *See* Don Peck, *They're Watching You at Work*, ATLANTIC MONTHLY, Dec. 2013, at 72, 76.

with points awarded when others use their code.[6] Individuals are assessed as likely to vote for a candidate based on their cable-usage patterns.[7] Recently released prisoners are scored on their likelihood of recidivism.[8]

How are these scores developed? Predictive algorithms mine personal information to make guesses about individuals' likely actions and risks.[9] A person's on- and offline activities are turned into scores that rate them above or below others.[10] Private and public entities rely on predictive algorithmic assessments to make important decisions about individuals.[11]

Sometimes, individuals can score the scorers, so to speak. Landlords can report bad tenants to data brokers while tenants can check abusive landlords on sites like ApartmentRatings.com. On sites like Rate My Professors, students can score professors who can respond to critiques via video. In many online communities, commenters can in turn rank the interplay between the rated, the raters, and the raters of the rated, in an effort to make sense of it all (or at least award the most convincing or popular with points or "karma").[12]

Although mutual-scoring opportunities among formally equal subjects exist in some communities, the realm of management and business more often features powerful entities who turn individuals into ranked and rated *objects*.[13] While scorers often characterize their work as an oasis of

[6] *See* E. GABRIELLA COLEMAN, CODING FREEDOM 116–22 (2013) (exploring Debian open source community and assessment of community members' contributions).

[7] *See* Alice E. Marwick, *How Your Data Are Being Deeply Mined*, N.Y. REV. BOOKS, Jan. 9, 2014, at 22, 22.

[8] Danielle Keats Citron, *Data Mining for Juvenile Offenders*, CONCURRING OPINIONS (Apr. 21, 2010, 3:56 PM), http://www.concurringopinions.com/archives/2010/04/data-mining-for-juvenile-offenders.html.

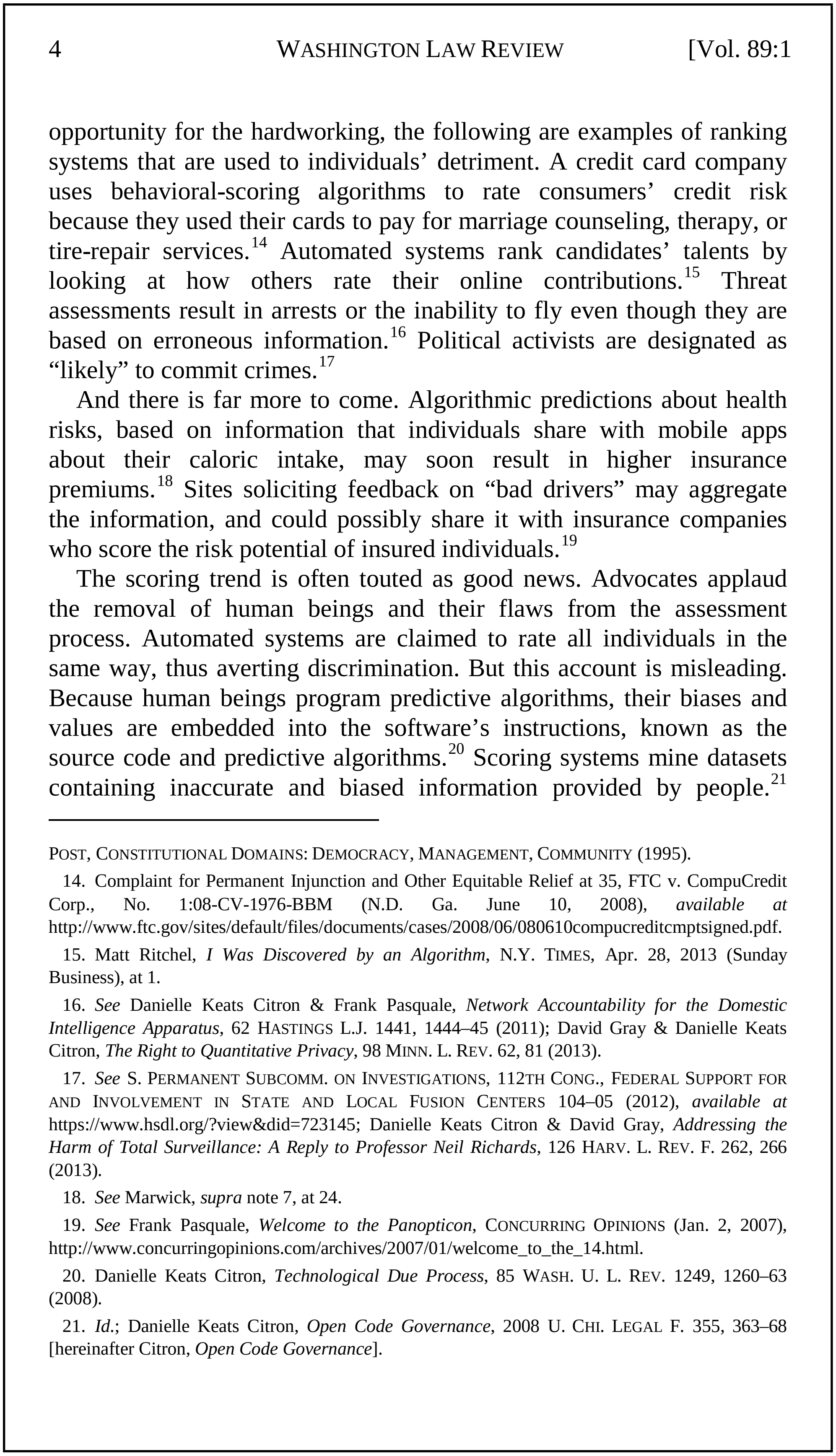
[9] Frank Pasquale, *Restoring Transparency to Automated Authority*, 9 J. ON TELECOMM. & HIGH TECH. L. 235, 235–36 (2011).

[10] Hussein A. Abdou & John Pointon, *Credit Scoring, Statistical Techniques and Evaluation Criteria: A Review of the Literature*, 18 INTELLIGENT SYSTEMS ACCT. FIN. & MGMT. 59, 60–61 (2011).

[11] *See* Marwick, *supra* note 7, at 24; *see also* Jack Nicas, *How Airlines Are Mining Personal Data In-Flight*, WALL ST. J., Nov. 8, 2013, at B1.

[12] Oren Bracha & Frank Pasquale, *Federal Search Commission? Access, Fairness, and Accountability in the Law of Search*, 93 CORNELL L. REV. 1149, 1159 (2008) ("This structures [sic] results in a bottom-up filtration system. At the lowest level, a large number of speakers receive relatively broad exposure within local communities likely composed of individuals with high-intensity interest or expertise. Speakers who gain salience at the lower levels may gradually gain recognition in higher-order clusters and eventually reach general visibility." (footnotes omitted)).

[13] *See* JARON LANIER, WHO OWNS THE FUTURE? 108 (2014); JARON LANIER, YOU ARE NOT A GADGET (2010). For the distinction between management and community, see generally ROBERT



opportunity for the hardworking, the following are examples of ranking systems that are used to individuals' detriment. A credit card company uses behavioral-scoring algorithms to rate consumers' credit risk because they used their cards to pay for marriage counseling, therapy, or tire-repair services.[14] Automated systems rank candidates' talents by looking at how others rate their online contributions.[15] Threat assessments result in arrests or the inability to fly even though they are based on erroneous information.[16] Political activists are designated as "likely" to commit crimes.[17]

And there is far more to come. Algorithmic predictions about health risks, based on information that individuals share with mobile apps about their caloric intake, may soon result in higher insurance premiums.[18] Sites soliciting feedback on "bad drivers" may aggregate the information, and could possibly share it with insurance companies who score the risk potential of insured individuals.[19]

The scoring trend is often touted as good news. Advocates applaud the removal of human beings and their flaws from the assessment process. Automated systems are claimed to rate all individuals in the same way, thus averting discrimination. But this account is misleading. Because human beings program predictive algorithms, their biases and values are embedded into the software's instructions, known as the source code and predictive algorithms.[20] Scoring systems mine datasets containing inaccurate and biased information provided by people.[21]

POST, CONSTITUTIONAL DOMAINS: DEMOCRACY, MANAGEMENT, COMMUNITY (1995).

[14] Complaint for Permanent Injunction and Other Equitable Relief at 35, FTC v. CompuCredit Corp., No. 1:08-CV-1976-BBM (N.D. Ga. June 10, 2008), *available at* http://www.ftc.gov/sites/default/files/documents/cases/2008/06/080610compucreditcmptsigned.pdf.

[15] Matt Ritchel, *I Was Discovered by an Algorithm*, N.Y. TIMES, Apr. 28, 2013 (Sunday Business), at 1.

[16] *See* Danielle Keats Citron & Frank Pasquale, *Network Accountability for the Domestic Intelligence Apparatus*, 62 HASTINGS L.J. 1441, 1444–45 (2011); David Gray & Danielle Keats Citron, *The Right to Quantitative Privacy*, 98 MINN. L. REV. 62, 81 (2013).

[17] *See* S. PERMANENT SUBCOMM. ON INVESTIGATIONS, 112TH CONG., FEDERAL SUPPORT FOR AND INVOLVEMENT IN STATE AND LOCAL FUSION CENTERS 104–05 (2012), *available at* https://www.hsdl.org/?view&did=723145; Danielle Keats Citron & David Gray, *Addressing the Harm of Total Surveillance: A Reply to Professor Neil Richards*, 126 HARV. L. REV. F. 262, 266 (2013).

[18] *See* Marwick, *supra* note 7, at 24.

[19] *See* Frank Pasquale, *Welcome to the Panopticon*, CONCURRING OPINIONS (Jan. 2, 2007), http://www.concurringopinions.com/archives/2007/01/welcome_to_the_14.html.

[20] Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1260–63 (2008).

[21] *Id.*; Danielle Keats Citron, *Open Code Governance*, 2008 U. CHI. LEGAL F. 355, 363–68 [hereinafter Citron, *Open Code Governance*].

There is nothing unbiased about scoring systems.

Supporters of scoring systems insist that we can trust algorithms to adjust themselves for greater accuracy. In the case of credit scoring, lenders combine the traditional three-digit credit scores with "credit analytics," which track consumers' transactions. Suppose credit-analytics systems predict that efforts to save money correlates with financial distress. Buying generic products instead of branded ones could then result in a hike in interest rates. But, the story goes, if consumers who bought generic brands also purchased items suggesting their financial strength, then all of their purchases would factor into their score, keeping them from being penalized from any particular purchase.

Does everything work out in a wash because information is seen in its totality? We cannot rigorously test this claim because scoring systems are shrouded in secrecy. Although some scores, such as credit, are available to the public, the scorers refuse to reveal the method and logic of their predictive systems.[22] No one can challenge the process of scoring and the results because the algorithms are zealously guarded trade secrets.[23] As this Article explores, the outputs of credit-scoring systems undermine supporters' claims. Credit scores are plagued by arbitrary results. They may also have a disparate impact on historically subordinated groups.

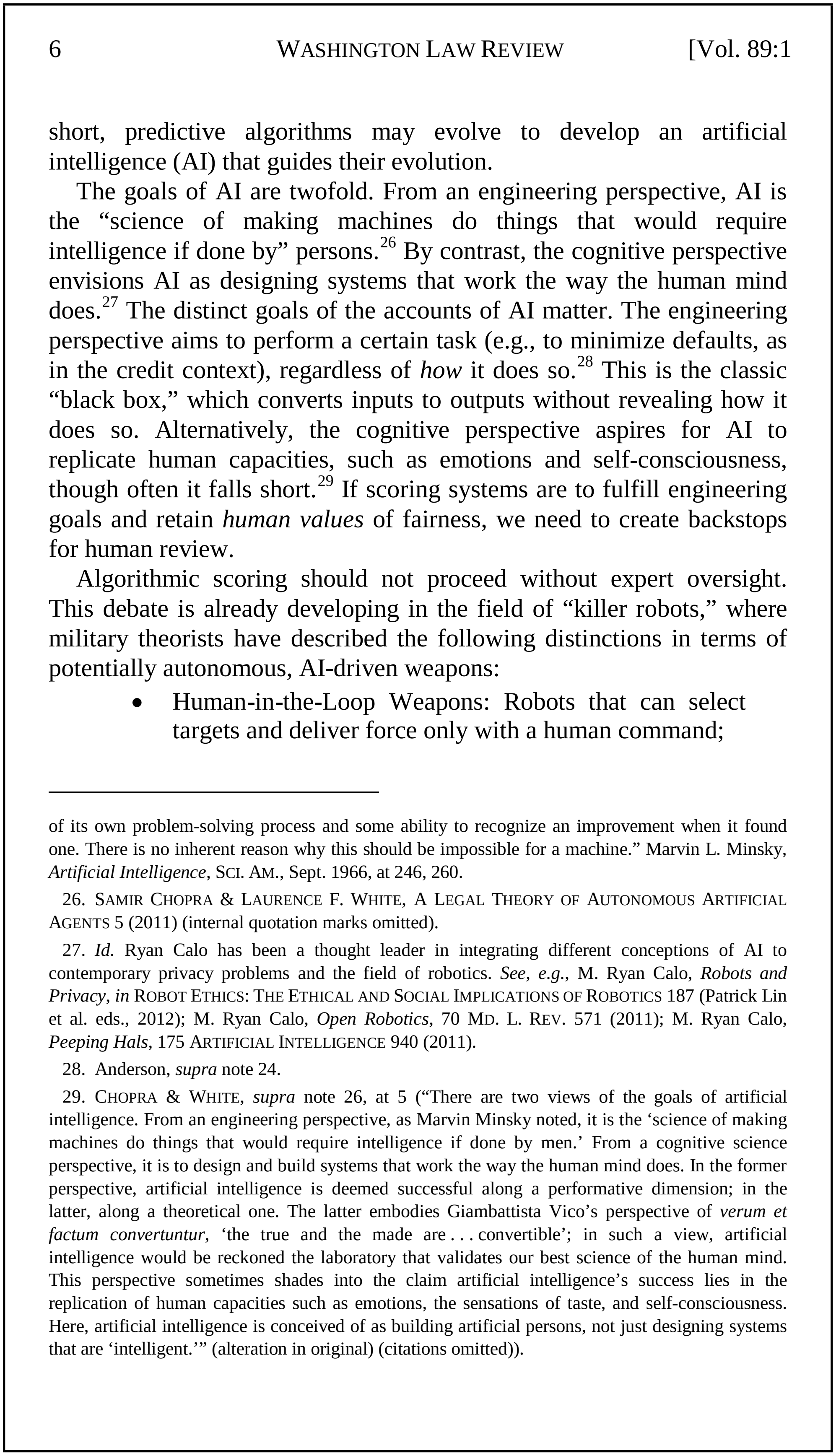
Just as concerns about scoring systems are more acute, their human element is diminishing. Although software engineers initially identify the correlations and inferences programmed into algorithms, Big Data promises to eliminate the human "middleman" at some point in the process.[24] Once data-mining programs have a range of correlations and inferences, they use them to project new forms of learning. The results of prior rounds of data mining can lead to unexpected correlations in click-through activity. If, for instance, predictive algorithms determine not only the types of behavior suggesting loan repayment, but also automate the process of learning which adjustments worked best in the past, the computing process reaches a third level of sophistication: determining *which* metrics for measuring past predictive algorithms were effective, and recommending further iterations for testing.[25] In

[22] Tal Zarsky, *Transparent Predictions*, 2013 ILL. L. REV. 1503, 1512.

[23] Evan Hendricks, *Credit Reports, Credit Checks, Credit Scores*, A.B.A. GPSOLO, July/Aug. 2011, at 32, 34.

[24] Chris Anderson, *The End of Theory: The Data Deluge Makes Scientific Inquiry Obsolete*, WIRED (June 23, 2008), http://www.wired.com/science/discoveries/magazine/16-07/pb_theory.

[25] A pioneer of artificial intelligence described this process in more general terms: "In order for a program to improve itself substantially it would have to have at least a rudimentary understanding



short, predictive algorithms may evolve to develop an artificial intelligence (AI) that guides their evolution.

The goals of AI are twofold. From an engineering perspective, AI is the "science of making machines do things that would require intelligence if done by" persons.[26] By contrast, the cognitive perspective envisions AI as designing systems that work the way the human mind does.[27] The distinct goals of the accounts of AI matter. The engineering perspective aims to perform a certain task (e.g., to minimize defaults, as in the credit context), regardless of *how* it does so.[28] This is the classic "black box," which converts inputs to outputs without revealing how it does so. Alternatively, the cognitive perspective aspires for AI to replicate human capacities, such as emotions and self-consciousness, though often it falls short.[29] If scoring systems are to fulfill engineering goals and retain *human values* of fairness, we need to create backstops for human review.

Algorithmic scoring should not proceed without expert oversight. This debate is already developing in the field of "killer robots," where military theorists have described the following distinctions in terms of potentially autonomous, AI-driven weapons:

- Human-in-the-Loop Weapons: Robots that can select targets and deliver force only with a human command;

of its own problem-solving process and some ability to recognize an improvement when it found one. There is no inherent reason why this should be impossible for a machine." Marvin L. Minsky, *Artificial Intelligence*, SCI. AM., Sept. 1966, at 246, 260.

[26] SAMIR CHOPRA & LAURENCE F. WHITE, A LEGAL THEORY OF AUTONOMOUS ARTIFICIAL AGENTS 5 (2011) (internal quotation marks omitted).

[27] *Id.* Ryan Calo has been a thought leader in integrating different conceptions of AI to contemporary privacy problems and the field of robotics. *See, e.g.*, M. Ryan Calo, *Robots and Privacy*, *in* ROBOT ETHICS: THE ETHICAL AND SOCIAL IMPLICATIONS OF ROBOTICS 187 (Patrick Lin et al. eds., 2012); M. Ryan Calo, *Open Robotics*, 70 MD. L. REV. 571 (2011); M. Ryan Calo, *Peeping Hals*, 175 ARTIFICIAL INTELLIGENCE 940 (2011).

[28] Anderson, *supra* note 24.

[29] CHOPRA & WHITE, *supra* note 26, at 5 ("There are two views of the goals of artificial intelligence. From an engineering perspective, as Marvin Minsky noted, it is the 'science of making machines do things that would require intelligence if done by men.' From a cognitive science perspective, it is to design and build systems that work the way the human mind does. In the former perspective, artificial intelligence is deemed successful along a performative dimension; in the latter, along a theoretical one. The latter embodies Giambattista Vico's perspective of *verum et factum convertuntur*, 'the true and the made are . . . convertible'; in such a view, artificial intelligence would be reckoned the laboratory that validates our best science of the human mind. This perspective sometimes shades into the claim artificial intelligence's success lies in the replication of human capacities such as emotions, the sensations of taste, and self-consciousness. Here, artificial intelligence is conceived of as building artificial persons, not just designing systems that are 'intelligent.'" (alteration in original) (citations omitted)).

- Human-on-the-Loop Weapons: Robots that can select targets and deliver force under the oversight of a human operator who can override the robots' actions; and
- Human-out-of-the-Loop Weapons: Robots that are capable of selecting targets and delivering force without any human input or interaction.[30]

Human rights advocates and computer scientists contend that "Human-out-of-the-Loop Weapons" systems violate international law because AI systems cannot adequately incorporate the rules of distinction ("which requires armed forces to distinguish between combatants and noncombatants") and proportionality.[31] They create a "responsibility gap" between commanders and killing machines.[32] Such decisions arguably are the unique responsibility of persons using holistic, non-algorithmic judgment to oversee complex and difficult situations.[33]

Just as automated killing machines violate basic legal norms, stigmatizing scoring systems at the least should be viewed with caution. We should not simply accept their predictions without understanding how they came about, and assuring that some human reviewer can respond to serious concerns about their fairness or accuracy.

Scoring systems are often assessed from an engineering perspective, as a calculative risk management technology making tough but ultimately technical rankings of populations as a whole. We call for the integration of the cognitive perspective of AI. In this Article, we explore the consequences to human values of fairness and justice when scoring machines make judgments about individuals. Although algorithmic predictions harm individuals' life opportunities often in arbitrary and discriminatory ways, they remain secret.[34] Human oversight is needed to

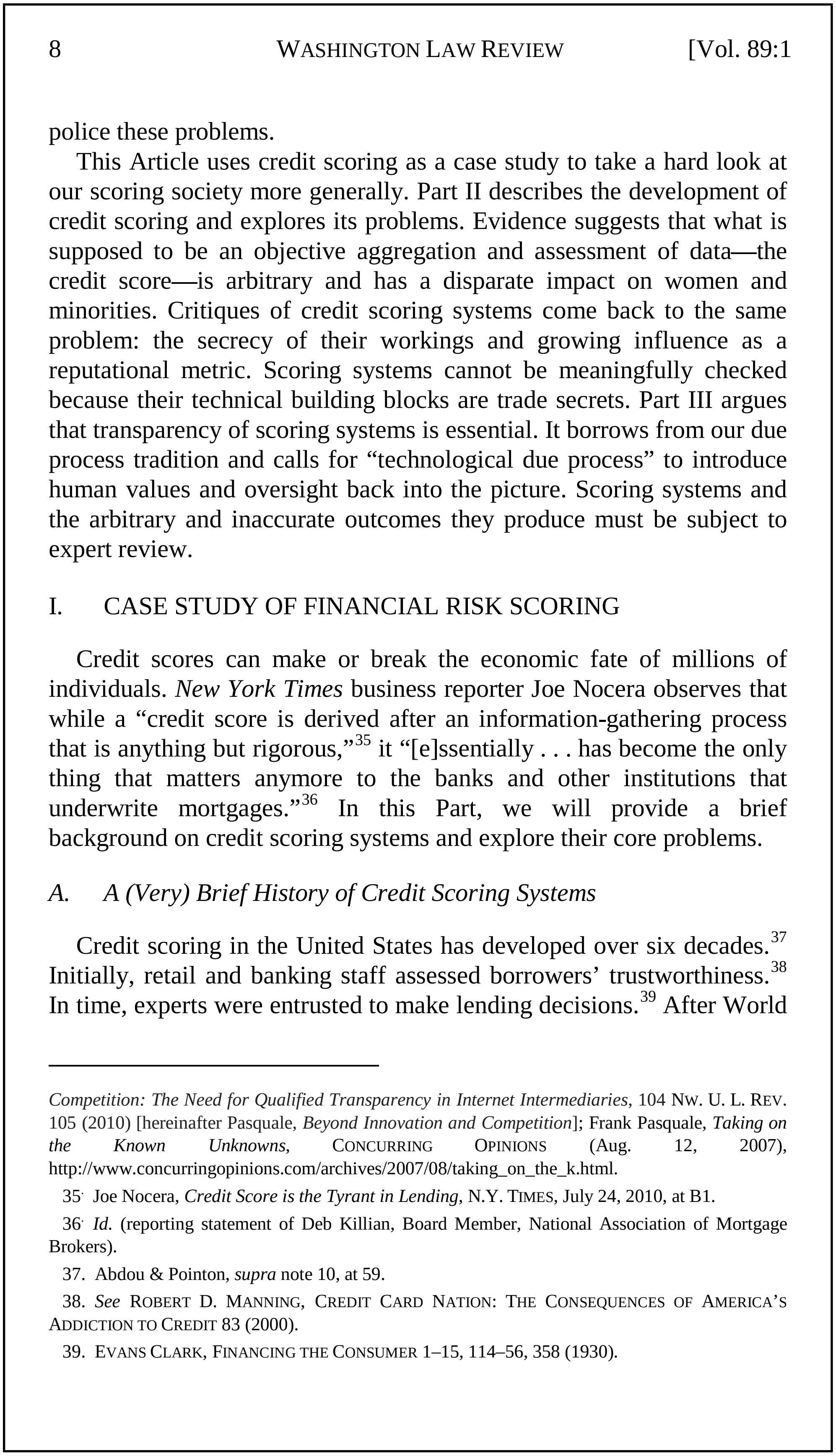
[30] *See Losing Humanity: The Case Against Killer Robots*, HUM. RTS. WATCH (Int'l Human Rights Clinic, Harvard Law Sch., Cambridge, Mass.), Nov. 2012, at 2, *available at* http://www.hrw.org/sites/default/files/reports/arms1112_ForUpload.pdf.

[31] *Id.* at 30.

[32] *Id.* at 42.

[33] *See* JOSEPH WEIZENBAUM, COMPUTER POWER AND HUMAN REASON: FROM JUDGMENT TO CALCULATION 227 (1976) (insisting that we should not delegate to computers "tasks that demand wisdom"). This is not to overstate the analogy of a low credit score to the kind of liberty deprivation at stake in weaponry. The *stakes* of war are far greater than being sure that an individual can be charged a higher interest rate. Nonetheless, under the *Mathews v. Eldridge*, 424 U.S. 319 (1976), calculus familiar to all students of administrative and constitutional law, *id.* at 332–39, we should not reject the targeting analogy as more-and-more predictive algorithms impact more-and-more aspects of our lives.

[34] On the importance of transparency and accountability in algorithms of powerful internet intermediaries, see Bracha & Pasquale, *supra* note 12; Frank Pasquale, *Beyond Innovation and*



police these problems.

This Article uses credit scoring as a case study to take a hard look at our scoring society more generally. Part II describes the development of credit scoring and explores its problems. Evidence suggests that what is supposed to be an objective aggregation and assessment of data—the credit score—is arbitrary and has a disparate impact on women and minorities. Critiques of credit scoring systems come back to the same problem: the secrecy of their workings and growing influence as a reputational metric. Scoring systems cannot be meaningfully checked because their technical building blocks are trade secrets. Part III argues that transparency of scoring systems is essential. It borrows from our due process tradition and calls for "technological due process" to introduce human values and oversight back into the picture. Scoring systems and the arbitrary and inaccurate outcomes they produce must be subject to expert review.

I. CASE STUDY OF FINANCIAL RISK SCORING

Credit scores can make or break the economic fate of millions of individuals. *New York Times* business reporter Joe Nocera observes that while a "credit score is derived after an information-gathering process that is anything but rigorous,"[35] it "[e]ssentially . . . has become the only thing that matters anymore to the banks and other institutions that underwrite mortgages."[36] In this Part, we will provide a brief background on credit scoring systems and explore their core problems.

A. A (Very) Brief History of Credit Scoring Systems

Credit scoring in the United States has developed over six decades.[37] Initially, retail and banking staff assessed borrowers' trustworthiness.[38] In time, experts were entrusted to make lending decisions.[39] After World

Competition: The Need for Qualified Transparency in Internet Intermediaries, 104 NW. U. L. REV. 105 (2010) [hereinafter Pasquale, *Beyond Innovation and Competition*]; Frank Pasquale, *Taking on the Known Unknowns*, CONCURRING OPINIONS (Aug. 12, 2007), http://www.concurringopinions.com/archives/2007/08/taking_on_the_k.html.

[35] Joe Nocera, *Credit Score is the Tyrant in Lending*, N.Y. TIMES, July 24, 2010, at B1.

[36] *Id.* (reporting statement of Deb Killian, Board Member, National Association of Mortgage Brokers).

[37] Abdou & Pointon, *supra* note 10, at 59.

[38] *See* ROBERT D. MANNING, CREDIT CARD NATION: THE CONSEQUENCES OF AMERICA'S ADDICTION TO CREDIT 83 (2000).

[39] EVANS CLARK, FINANCING THE CONSUMER 1–15, 114–56, 358 (1930).

War II, specialized finance companies entered the mix.[40]

In 1956, the firm Fair, Isaac & Co. (now known as FICO) devised a three-digit credit score, promoting its services to banks and finance companies.[41] FICO marketed its scores as predictors of whether consumers would default on their debts.[42] FICO scores range from 300 to 850. FICO's scoring system remains powerful, though credit bureaus ("consumer reporting agencies")[43] have developed their own scoring systems as well.[44]

Credit scores legitimated the complex securities at the heart of the recent financial crisis.[45] In the mid-2000s, the credit score was the key connecting ordinary U.S. homeowners with international capital investors eager to invest in highly rated securities.[46] When investors purchased a mortgage-backed security, they bought the right to a stream of payments.[47] The mortgagor (borrower) shifted from paying the

[40] Stan Sienkiewicz, Credit Cards and Payment Efficiency 3 (Aug. 2001) (unpublished discussion paper), *available at* http://www.philadelphiafed.org/consumer-credit-and-payments/payment-cards-center/publications/discussionpapers/2001/PaymentEfficiency_092001.pdf.

[41] Martha Poon, *Scorecards as Market Devices for Consumer Credit: The Case of Fair, Isaac & Company Incorporated*, 55 SOC. REV. MONOGRAPH 284, 288 (2007); *Fair, Isaac and Company History*, FUNDINGUNIVERSE, http://www.fundinguniverse.com/company-histories/Fair-Isaac-and-Company-Company-History.html (last visited Feb. 8, 2014).

[42] On predicting "derogatory events," see *The FICO Score*, THECREDITSCORINGSITE, http://www.creditscoring.com/creditscore/fico/ (last visited Feb. 8, 2014).

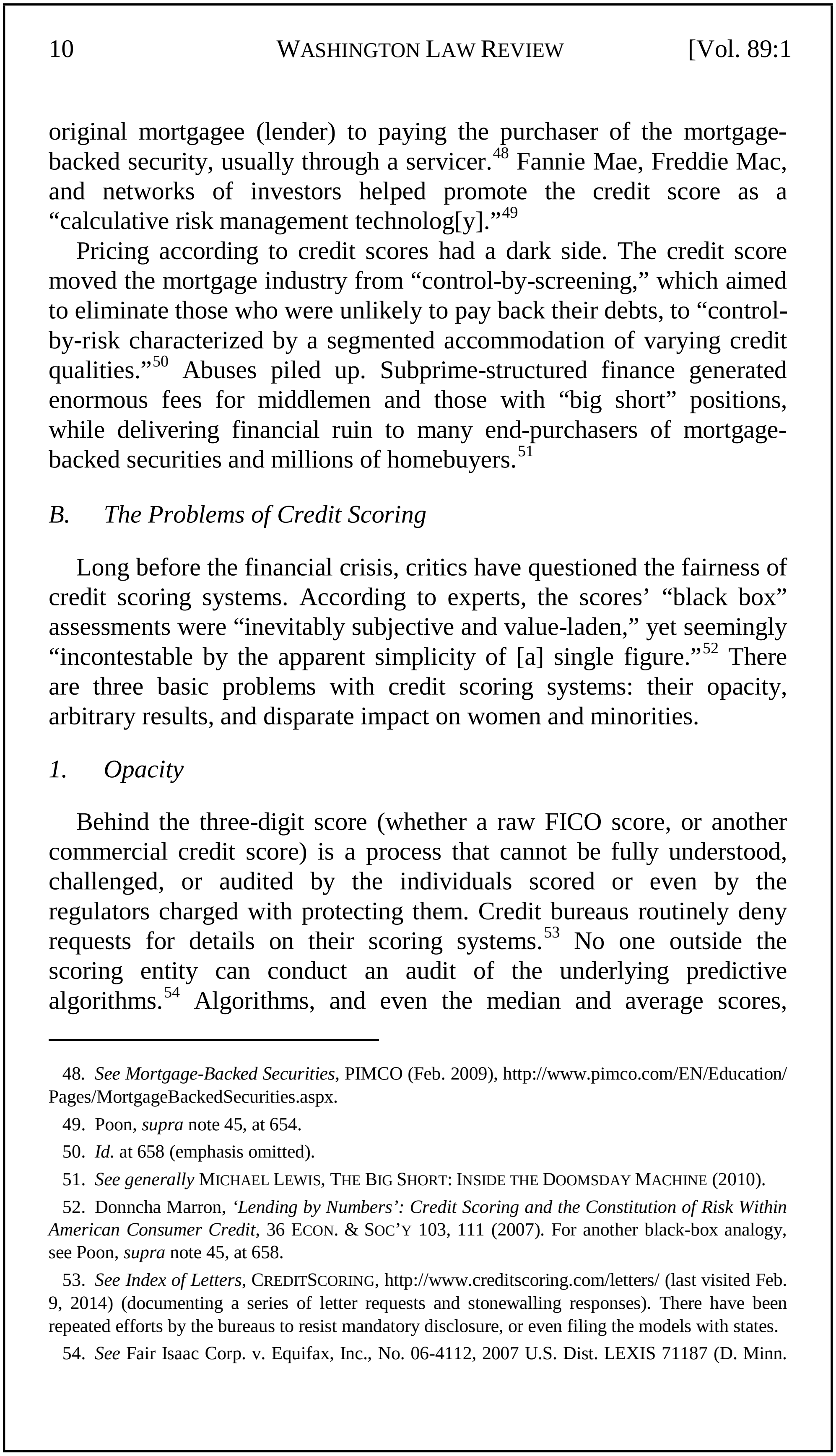
[43] For a definition of credit bureau, see *Elkins v. Ocwen Federal Savings Bank Experian Information Solutions, Inc.*, No. 06 CV 823, 2007 U.S. Dist. LEXIS 84556, at *36–37 (N.D. Ill. Nov. 13, 2007) (explaining that credit bureaus and consumer reporting agencies regularly receive updates on a consumer's credit relationships from their data furnishers, such as banks, mortgage companies, debt collectors, credit card issuers, department stores and others, and produce reports that contain highly sensitive and personal details about a consumer's finances, including account numbers, loan balances, credit limits, and payment history).

[44] A court case describes the fight between FICO and credit bureaus over the credit bureaus' development of their own scoring systems. *See* Fair Isaac Corp. v. Experian Info. Solutions, Inc., 645 F. Supp. 2d 734 (D. Minn. 2009), *aff'd*, 650 F.3d 1139 (8th Cir. 2011). In such cases, courts use protective orders to ensure the confidentiality of trade secrets. *See, e.g.*, Textured Yarn Co. v. Burkart-Schier Chem. Co., 41 F.R.D. 158 (E.D. Tenn. 1966).

[45] Martha Poon, *From New Deal Institutions to Capital Markets: Commercial Consumer Risk Scores and the Making of Subprime Mortgage Finance*, 34 ACCT. ORGS. & SOC'Y, 654, 662 (2009). In 1995, government-sponsored entities Fannie Mae and Freddie Mac announced that borrowers needed a credit score of at least 660 (on FICO's scale of 300 to 850) for loans to qualify for the status of "prime investment." *Id.* at 663. Those below 660 were relegated to "subprime" offerings. *Id.* at 664.

[46] *Id.* at 655.

[47] Chris Wilson, *What Is a Mortgage-Backed Security?*, SLATE (Mar. 17, 2008, 7:09 PM), http://www.slate.com/articles/news_and_politics/explainer/2008/03/what_is_a_mortgagebacked_security.html.



original mortgagee (lender) to paying the purchaser of the mortgage-backed security, usually through a servicer.[48] Fannie Mae, Freddie Mac, and networks of investors helped promote the credit score as a "calculative risk management technolog[y]."[49]

Pricing according to credit scores had a dark side. The credit score moved the mortgage industry from "control-by-screening," which aimed to eliminate those who were unlikely to pay back their debts, to "control-by-risk characterized by a segmented accommodation of varying credit qualities."[50] Abuses piled up. Subprime-structured finance generated enormous fees for middlemen and those with "big short" positions, while delivering financial ruin to many end-purchasers of mortgage-backed securities and millions of homebuyers.[51]

B. *The Problems of Credit Scoring*

Long before the financial crisis, critics have questioned the fairness of credit scoring systems. According to experts, the scores' "black box" assessments were "inevitably subjective and value-laden," yet seemingly "incontestable by the apparent simplicity of [a] single figure."[52] There are three basic problems with credit scoring systems: their opacity, arbitrary results, and disparate impact on women and minorities.

1. *Opacity*

Behind the three-digit score (whether a raw FICO score, or another commercial credit score) is a process that cannot be fully understood, challenged, or audited by the individuals scored or even by the regulators charged with protecting them. Credit bureaus routinely deny requests for details on their scoring systems.[53] No one outside the scoring entity can conduct an audit of the underlying predictive algorithms.[54] Algorithms, and even the median and average scores,

[48] *See Mortgage-Backed Securities*, PIMCO (Feb. 2009), http://www.pimco.com/EN/Education/Pages/MortgageBackedSecurities.aspx.

[49] Poon, *supra* note 45, at 654.

[50] *Id.* at 658 (emphasis omitted).

[51] *See generally* MICHAEL LEWIS, THE BIG SHORT: INSIDE THE DOOMSDAY MACHINE (2010).

[52] Donncha Marron, *'Lending by Numbers': Credit Scoring and the Constitution of Risk Within American Consumer Credit*, 36 ECON. & SOC'Y 103, 111 (2007). For another black-box analogy, see Poon, *supra* note 45, at 658.

[53] *See Index of Letters*, CREDITSCORING, http://www.creditscoring.com/letters/ (last visited Feb. 9, 2014) (documenting a series of letter requests and stonewalling responses). There have been repeated efforts by the bureaus to resist mandatory disclosure, or even filing the models with states.

[54] *See* Fair Isaac Corp. v. Equifax, Inc., No. 06-4112, 2007 U.S. Dist. LEXIS 71187 (D. Minn.

remain secret.

The lack of transparency of credit-scoring systems leaves consumers confounded by how and why their scores change.[55] FICO and the credit bureaus do not explain the extent to which individual behavior affects certain categories.[56] Consumers cannot determine optimal credit behavior or even what to do to avoid a hit on their scores.

FICO and credit bureaus do, however, announce the relative weight of certain categories in their scoring systems.[57] For example, "credit utilization" (how much of a borrower's current credit lines are being used) may be used. But the optimal credit utilization strategy is unclear. No one knows whether, for instance, using twenty-five percent of one's credit limit is better or worse than using fifteen percent. An ambitious consumer could try to reverse-engineer credit scores, but such efforts would be expensive and unreliable.[58]

As various rankings proliferate, so do uncertainties about one's standing.[59] Even the most conscientious borrower may end up surprised by the consequences of his actions. Responding to the confusion, books, articles, and websites offer advice on scoring systems. Amazon offers dozens of self-help books on the topic, each capitalizing on credit scoring's simultaneously mystifying and meritocratic reputation.[60] Hucksters abound in the cottage industry of do-it-yourself credit repair.

2. *Arbitrary Assessments*

Credit-scoring systems produce arbitrary results, as demonstrated by

Sept. 25, 2007); *see also* Public Comment Letter from Greg Fisher, Creditscoring.com, to the Board of Governors of the Federal Reserve (Sept. 17, 2004), *available at* http://www.federalreserve.gov/SECRS/2004/October/20041014/OP-1209/OP-1209_106_1.pdf.

[55] Yuliya Demyanyk, *Your Credit Score Is a Ranking, Not a Score*, FED. RES. BANK CLEVELAND (Nov. 16, 2010), http://www.clevelandfed.org/research/commentary/2010/2010-16.cfm.

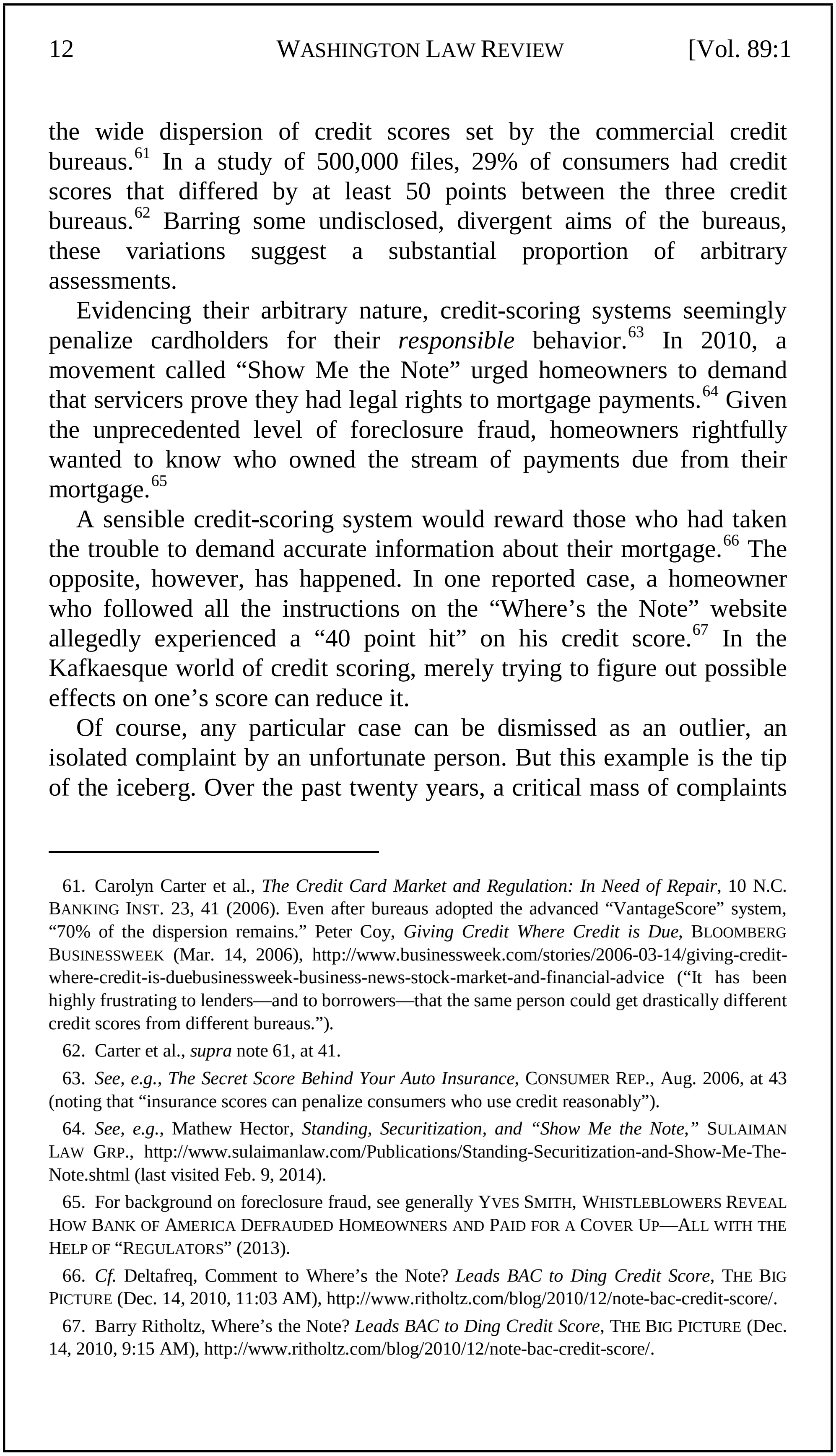
[56] *Credit Checks & Inquiries*, MYFICO, http://www.myfico.com/crediteducation/creditinquiries.aspx (last visited Feb. 9, 2014).

[57] *See, e.g.*, *What's in My FICO Score?*, MYFICO, http://www.myfico.com/CreditEducation/WhatsInYourScore.aspx (last visited Feb. 9, 2014).

[58] *See* Dean Foust & Aaron Pressman, *Credit Scores: Not-So-Magic Numbers*, BLOOMBERG BUSINESSWEEK (Feb. 6, 2008), http://www.businessweek.com/stories/2008-02-06/credit-scores-not-so-magic-numbers.

[59] *See, e.g.*, Sue Kirchhoff & Sandra Block, *Alternative Credit Scores Could Open Door for Loans*, USA TODAY (May 16, 2006, 10:01 PM), http://usatoday30.usatoday.com/money/perfi/credit/2006-05-16-credit-scores-usat_x.htm.

[60] *See, e.g., Owing! 5 Lessons on Surviving Your Debt Living in a Culture of Credit [Kindle Edition]*, AMAZON.COM, http://www.amazon.com/Lessons-Surviving-Living-Culture-Credit-ebook/dp/B00C2BMN3W (last visited Feb. 12, 2014).



the wide dispersion of credit scores set by the commercial credit bureaus.[61] In a study of 500,000 files, 29% of consumers had credit scores that differed by at least 50 points between the three credit bureaus.[62] Barring some undisclosed, divergent aims of the bureaus, these variations suggest a substantial proportion of arbitrary assessments.

Evidencing their arbitrary nature, credit-scoring systems seemingly penalize cardholders for their *responsible* behavior.[63] In 2010, a movement called "Show Me the Note" urged homeowners to demand that servicers prove they had legal rights to mortgage payments.[64] Given the unprecedented level of foreclosure fraud, homeowners rightfully wanted to know who owned the stream of payments due from their mortgage.[65]

A sensible credit-scoring system would reward those who had taken the trouble to demand accurate information about their mortgage.[66] The opposite, however, has happened. In one reported case, a homeowner who followed all the instructions on the "Where's the Note" website allegedly experienced a "40 point hit" on his credit score.[67] In the Kafkaesque world of credit scoring, merely trying to figure out possible effects on one's score can reduce it.

Of course, any particular case can be dismissed as an outlier, an isolated complaint by an unfortunate person. But this example is the tip of the iceberg. Over the past twenty years, a critical mass of complaints

[61] Carolyn Carter et al., *The Credit Card Market and Regulation: In Need of Repair*, 10 N.C. BANKING INST. 23, 41 (2006). Even after bureaus adopted the advanced "VantageScore" system, "70% of the dispersion remains." Peter Coy, *Giving Credit Where Credit is Due*, BLOOMBERG BUSINESSWEEK (Mar. 14, 2006), http://www.businessweek.com/stories/2006-03-14/giving-credit-where-credit-is-duebusinessweek-business-news-stock-market-and-financial-advice ("It has been highly frustrating to lenders—and to borrowers—that the same person could get drastically different credit scores from different bureaus.").

[62] Carter et al., *supra* note 61, at 41.

[63] *See, e.g.*, *The Secret Score Behind Your Auto Insurance*, CONSUMER REP., Aug. 2006, at 43 (noting that "insurance scores can penalize consumers who use credit reasonably").

[64] *See, e.g.*, Mathew Hector, *Standing, Securitization, and "Show Me the Note,"* SULAIMAN LAW GRP., http://www.sulaimanlaw.com/Publications/Standing-Securitization-and-Show-Me-The-Note.shtml (last visited Feb. 9, 2014).

[65] For background on foreclosure fraud, see generally YVES SMITH, WHISTLEBLOWERS REVEAL HOW BANK OF AMERICA DEFRAUDED HOMEOWNERS AND PAID FOR A COVER UP—ALL WITH THE HELP OF "REGULATORS" (2013).

[66] *Cf.* Deltafreq, Comment to Where's the Note? *Leads BAC to Ding Credit Score*, THE BIG PICTURE (Dec. 14, 2010, 11:03 AM), http://www.ritholtz.com/blog/2010/12/note-bac-credit-score/.

[67] Barry Ritholtz, Where's the Note? *Leads BAC to Ding Credit Score*, THE BIG PICTURE (Dec. 14, 2010, 9:15 AM), http://www.ritholtz.com/blog/2010/12/note-bac-credit-score/.

about credit scoring has emerged.[68] Cassandra Jones Havard contends that scoring models may play an integral role in discriminatory lending practices.[69] Another commentator has charged that they enabled reckless securitizations that had devastating systemic impact.[70]

In many accounts of the financial crisis, credit scores exerted a baleful influence, rationalizing lending practices with ersatz quantification. As Amar Bhide argued, the idea of "one best way" to rank credit applicants flattened the distributed, varying judgment of local loan officers into the nationwide credit score—a number focused on *persons* rather than *communities*.[71] Like monocultural-farming technology vulnerable to one unanticipated bug, the converging methods of credit assessment failed spectacularly when macroeconomic conditions changed. The illusion of commensurability and solid valuation provided by the models that mortgage-based securities were based on helped spark a rush for what appeared to be easy returns, exacerbating both boom and bust dynamics.

3. Disparate Impact

Far from eliminating existing discriminatory practices, credit-scoring algorithms instead grant them an imprimatur, systematizing them in hidden ways.[72] Credit scores are only as free from bias as the software

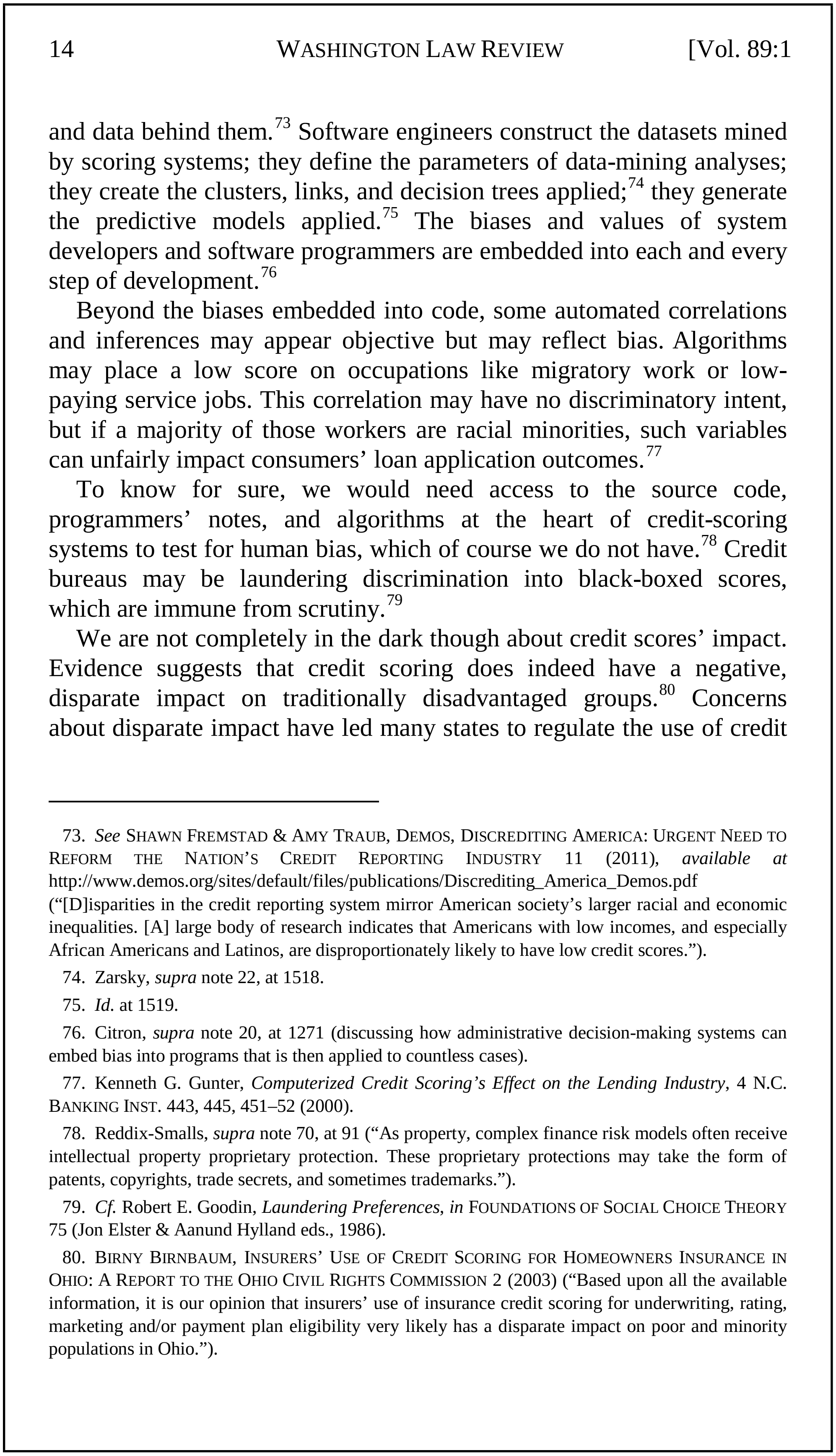
[68] *See, e.g.*, Kevin Simpson, *Insurers' Use of Credit Reports Rankles Many*, DENVER POST, Aug. 20, 2003, at A1 ("Credit-scoring has been one of the components responsible for an 'alarming trend' of increased complaints to regulators over the past three years . . . .").

[69] Cassandra Jones Havard, *"On The Take": The Black Box of Credit Scoring and Mortgage Discrimination*, 20 B.U. PUB. INT. L.J. 241, 247 (2011) (arguing that credit scoring if unchecked is an intrinsic, established form of discrimination very similar to redlining).

[70] Brenda Reddix-Smalls, *Credit Scoring and Trade Secrecy: An Algorithmic Quagmire or How the Lack of Transparency in Complex Financial Models Scuttled the Finance Market*, 12 U.C. DAVIS BUS. L.J. 87 (2011).

[71] *See generally* AMAR BHIDE, A CALL FOR JUDGMENT: SENSIBLE FINANCE FOR A DYNAMIC ECONOMY (2010); Meredith Schramm-Strosser, *The "Not So" Fair Credit Reporting Act: Federal Preemption, Injunctive Relief, and the Need to Return Remedies for Common Law Defamation to the States*, 14 DUQ. BUS. L.J. 165, 169 (2012) ("A consumer's reputation and credibility is determined not by personal interactions with others in a small community, but by examining credit files in an impersonal global world.").

[72] Havard, *supra* note 69, at 247 (arguing that "credit scoring if unchecked is an intrinsic, established form of discrimination very similar to redlining"). *Cf.* Citron & Pasquale, *supra* note 16, at 1459 (exploring how bias against groups can be embedded in fusion centers' data-mining algorithms and spread through the information sharing environment). The EEOC, in a lawsuit filed against Kaplan, claimed that use of credit history would have a disparate, negative impact against minority job applicants because of the lower average credit score of these groups. Press Release, Equal Emp't Opportunity Comm'n, EEOC Files Nationwide Hiring Discrimination Lawsuit Against Kaplan Higher Education Corp. (Dec. 21, 2010), *available at* http://www.eeoc.gov/eeoc/newsroom/release/12-21-10a.cfm.



and data behind them.[73] Software engineers construct the datasets mined by scoring systems; they define the parameters of data-mining analyses; they create the clusters, links, and decision trees applied;[74] they generate the predictive models applied.[75] The biases and values of system developers and software programmers are embedded into each and every step of development.[76]

Beyond the biases embedded into code, some automated correlations and inferences may appear objective but may reflect bias. Algorithms may place a low score on occupations like migratory work or low-paying service jobs. This correlation may have no discriminatory intent, but if a majority of those workers are racial minorities, such variables can unfairly impact consumers' loan application outcomes.[77]

To know for sure, we would need access to the source code, programmers' notes, and algorithms at the heart of credit-scoring systems to test for human bias, which of course we do not have.[78] Credit bureaus may be laundering discrimination into black-boxed scores, which are immune from scrutiny.[79]

We are not completely in the dark though about credit scores' impact. Evidence suggests that credit scoring does indeed have a negative, disparate impact on traditionally disadvantaged groups.[80] Concerns about disparate impact have led many states to regulate the use of credit

[73] *See* Shawn Fremstad & Amy Traub, Demos, Discrediting America: Urgent Need to Reform the Nation's Credit Reporting Industry 11 (2011), *available at* http://www.demos.org/sites/default/files/publications/Discrediting_America_Demos.pdf ("[D]isparities in the credit reporting system mirror American society's larger racial and economic inequalities. [A] large body of research indicates that Americans with low incomes, and especially African Americans and Latinos, are disproportionately likely to have low credit scores.").

[74] Zarsky, *supra* note 22, at 1518.

[75] *Id.* at 1519.

[76] Citron, *supra* note 20, at 1271 (discussing how administrative decision-making systems can embed bias into programs that is then applied to countless cases).

[77] Kenneth G. Gunter, *Computerized Credit Scoring's Effect on the Lending Industry*, 4 N.C. Banking Inst. 443, 445, 451–52 (2000).

[78] Reddix-Smalls, *supra* note 70, at 91 ("As property, complex finance risk models often receive intellectual property proprietary protection. These proprietary protections may take the form of patents, copyrights, trade secrets, and sometimes trademarks.").

[79] *Cf.* Robert E. Goodin, *Laundering Preferences*, *in* Foundations of Social Choice Theory 75 (Jon Elster & Aanund Hylland eds., 1986).

[80] Birny Birnbaum, Insurers' Use of Credit Scoring for Homeowners Insurance in Ohio: A Report to the Ohio Civil Rights Commission 2 (2003) ("Based upon all the available information, it is our opinion that insurers' use of insurance credit scoring for underwriting, rating, marketing and/or payment plan eligibility very likely has a disparate impact on poor and minority populations in Ohio.").

scores in insurance underwriting.[81] The National Fair Housing Alliance (NFHA) has criticized credit scores for disadvantaging women and minorities.[82]

Insurers' use of credit scores has been challenged in court for their disparate impact on minorities. After years of litigation, Allstate agreed to a multi-million dollar settlement over "deficiencies in Allstate's credit scoring procedure which plaintiffs say resulted in discriminatory action against approximately five million African-American and Hispanic customers."[83] As part of the settlement, Allstate allowed plaintiffs' experts to critique and refine future scoring models.[84]

If illegal or unethical discrimination influences credit scoring, members of disadvantaged groups will have difficulty paying their bills.[85] Their late payments could be fed into credit scoring models as neutral, objective indicia of reliability and creditworthiness.[86] The very benchmark against which discriminatory practices are measured may indeed be influenced by discriminatory practices.

The paucity of enforcement activity makes it hard to assess the effectiveness of the Equal Credit Opportunity Act (ECOA), which prohibits discrimination in lending, and Regulation B, which applies ECOA to credit scoring systems.[87] Regulation B requires that the reasons for a denial of credit/lending has to be related to—and

[81] *Credit-Based Insurance Scoring: Separating Facts from Fallacies*, NAMIC POL'Y BRIEFING (Nat'l Ass'n of Mut. Ins. Cos., Indianapolis, Ind.), Feb. 2010, at 1, *available at* http://iiky.org/documents/NAMIC_Policy_Briefing_on_Insurance_Scoring_Feb_2010.pdf.

[82] *The Future of Housing Finance: The Role of Private Mortgage Insurance: Hearing Before the Subcomm. on Capital Mkts., Ins. & Gov't Sponsored Enters. of the H. Comm. on Fin. Servs.*, 111th Cong. 16 (2010) (statement of Deborah Goldberg, Hurricane Relief Program Director, The National Fair Housing Alliance). The NFHA has expressed concern that "the use of credit scores tends to disadvantage people of color, women, and others whose scores are often lower than those of white borrowers." *Id.* at 57. The NFHA has also expressed "growing concern about how useful credit scores are for predicting loan performance and whether the financial sector is placing too much reliance on credit scores rather than other risk factors such as loan terms." *Id.*

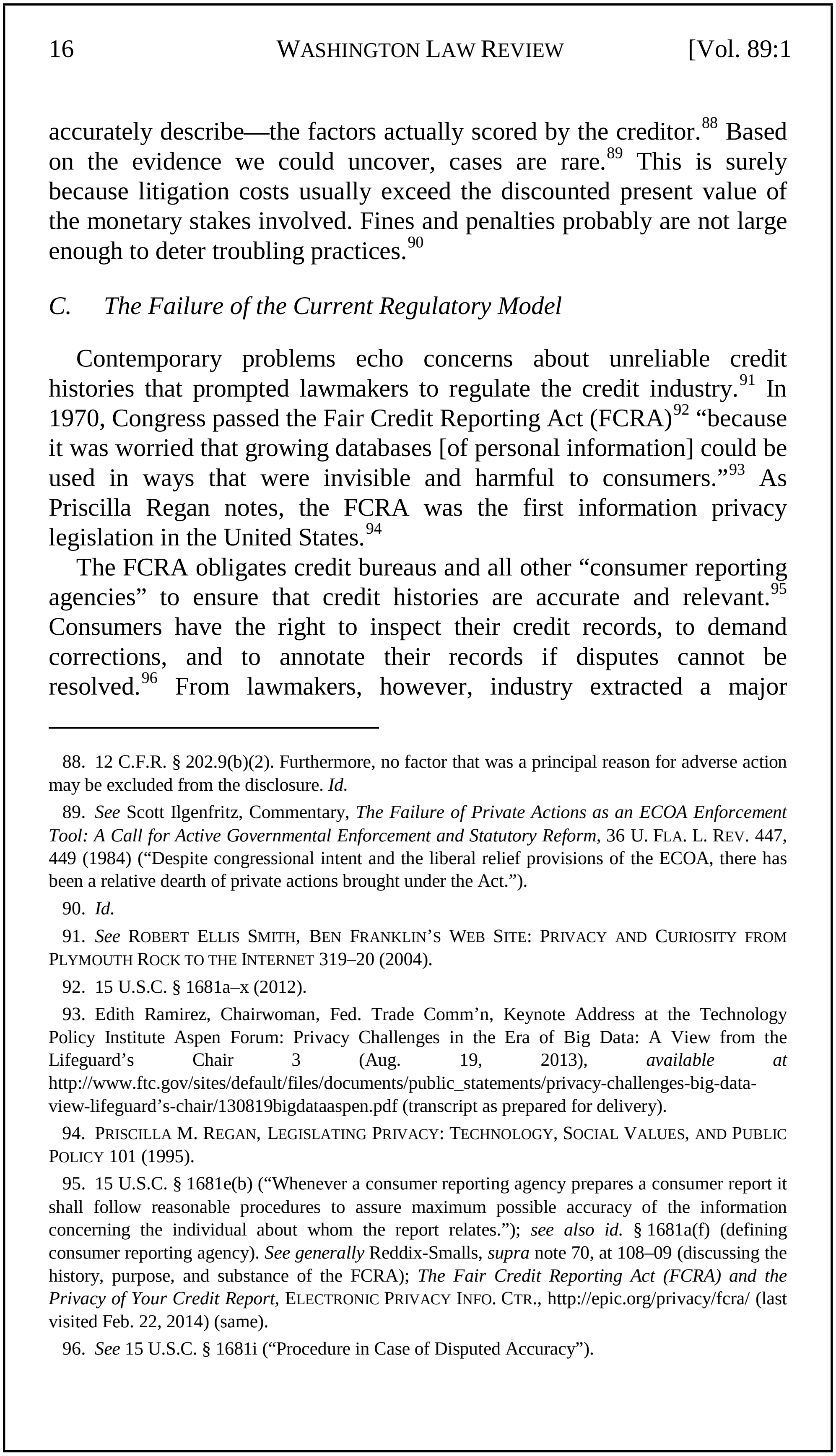
[83] Dehoyos v. Allstate, 240 F.R.D. 269, 275 (W.D. Tex. 2007). The parties settled after the Fifth Circuit decided that federal civil rights law was not reverse preempted by the McCarran-Ferguson Act's allocation of insurance regulatory authority to states. *See* Dehoyos v. Allstate Corp., 345 F.3d 290, 299 (5th Cir. 2003). The Equal Credit Opportunity Act (ECOA), which regulates lending practices, does not preempt state laws that are stricter than ECOA.

[84] *Dehoyos*, 240 F.R.D. at 276.

[85] *See, e.g.*, Gunter, *supra* note 77, at 451–52.

[86] *See generally* BERNARD E. HARCOURT, AGAINST PREDICTION: PROFILING, POLICING, AND PUNISHING IN AN ACTUARIAL AGE (2007).

[87] Regulation B sets forth specific data that cannot be used in a credit scoring system, such as: public assistance status, likelihood that any person will bear or rear children, telephone listing, income because of a prohibited basis, inaccurate credit histories, and different standards for married and unmarried persons, race, color, religion, national origin, and sex. 12 C.F.R. § 202.5 (2013).



accurately describe—the factors actually scored by the creditor.[88] Based on the evidence we could uncover, cases are rare.[89] This is surely because litigation costs usually exceed the discounted present value of the monetary stakes involved. Fines and penalties probably are not large enough to deter troubling practices.[90]

C. *The Failure of the Current Regulatory Model*

Contemporary problems echo concerns about unreliable credit histories that prompted lawmakers to regulate the credit industry.[91] In 1970, Congress passed the Fair Credit Reporting Act (FCRA)[92] "because it was worried that growing databases [of personal information] could be used in ways that were invisible and harmful to consumers."[93] As Priscilla Regan notes, the FCRA was the first information privacy legislation in the United States.[94]

The FCRA obligates credit bureaus and all other "consumer reporting agencies" to ensure that credit histories are accurate and relevant.[95] Consumers have the right to inspect their credit records, to demand corrections, and to annotate their records if disputes cannot be resolved.[96] From lawmakers, however, industry extracted a major

[88] 12 C.F.R. § 202.9(b)(2). Furthermore, no factor that was a principal reason for adverse action may be excluded from the disclosure. *Id.*

[89] *See* Scott Ilgenfritz, Commentary, *The Failure of Private Actions as an ECOA Enforcement Tool: A Call for Active Governmental Enforcement and Statutory Reform*, 36 U. FLA. L. REV. 447, 449 (1984) ("Despite congressional intent and the liberal relief provisions of the ECOA, there has been a relative dearth of private actions brought under the Act.").

[90] *Id.*

[91] *See* ROBERT ELLIS SMITH, BEN FRANKLIN'S WEB SITE: PRIVACY AND CURIOSITY FROM PLYMOUTH ROCK TO THE INTERNET 319–20 (2004).

[92] 15 U.S.C. § 1681a–x (2012).

[93] Edith Ramirez, Chairwoman, Fed. Trade Comm'n, Keynote Address at the Technology Policy Institute Aspen Forum: Privacy Challenges in the Era of Big Data: A View from the Lifeguard's Chair 3 (Aug. 19, 2013), *available at* http://www.ftc.gov/sites/default/files/documents/public_statements/privacy-challenges-big-data-view-lifeguard's-chair/130819bigdataaspen.pdf (transcript as prepared for delivery).

[94] PRISCILLA M. REGAN, LEGISLATING PRIVACY: TECHNOLOGY, SOCIAL VALUES, AND PUBLIC POLICY 101 (1995).

[95] 15 U.S.C. § 1681e(b) ("Whenever a consumer reporting agency prepares a consumer report it shall follow reasonable procedures to assure maximum possible accuracy of the information concerning the individual about whom the report relates."); *see also id.* § 1681a(f) (defining consumer reporting agency). *See generally* Reddix-Smalls, *supra* note 70, at 108–09 (discussing the history, purpose, and substance of the FCRA); *The Fair Credit Reporting Act (FCRA) and the Privacy of Your Credit Report*, ELECTRONIC PRIVACY INFO. CTR., http://epic.org/privacy/fcra/ (last visited Feb. 22, 2014) (same).

[96] *See* 15 U.S.C. § 1681i ("Procedure in Case of Disputed Accuracy").

concession: immunity from defamation law.[97] By limiting the possible penalties for reputational injuries, the FCRA opened the door to tactics of stalling, obstinacy, and obfuscation by the credit industry.[98]

What about credit scores? In 2003, the Fair and Accurate Credit Transactions Act (FACTA) required credit bureaus to disclose credit scores to individuals in exchange for a fee capped by the FTC. But the FACTA does *not* "require a consumer reporting agency to disclose to a consumer any information concerning credit scores or any other risk scores or predictors relating to the consumer,"[99] except for four "key factors" involved in credit decisions.[100] Regrettably, those four factors do little to explain credit scores. Phrases like "type of bank accounts" and "type of credit references" are etiolated symbols, more suited for machine-to-machine interaction than personal explanation. Factors such as "too many revolving accounts" and "late payment" are commonplace even for those with high credit scores.[101] The law does not require credit scorers to tell individuals *how much* any given factor mattered to a particular score.[102] Looking forward, a consumer has no idea, for example, whether paying off a debt that is sixty days past due will raise her score. The industry remains highly opaque, with scored individuals unable to determine the exact consequences of their decisions.

Although FCRA offers individuals a chance to dispute items on their credit *history*, it does not require credit bureaus to reveal the way they convert a history into a score.[103] That is a trade secret; a designation offering powerful legal protections to companies that want to keep their business practices a secret.[104] Despite such secrecy, we can draw some

[97] SMITH, *supra* note 91, at 320. Note, though, that the FCRA is riddled with many exceptions, exceptions to exceptions, and interactions with state law.

[98] *See* Schramm-Strosser, *supra* note 71, at 170–71 ("What started out as an improvement over how the common law dealt with credit-reporting issues has evolved into a regulatory scheme that tends to favor the credit reporting industry . . . . One example of the FCRA's overly broad preemptive scope is the prohibition of injunctive relief for consumers who bring common law defamation claims against CRAs.").

[99] 15 U.S.C. § 1681g(a)(1)(B).

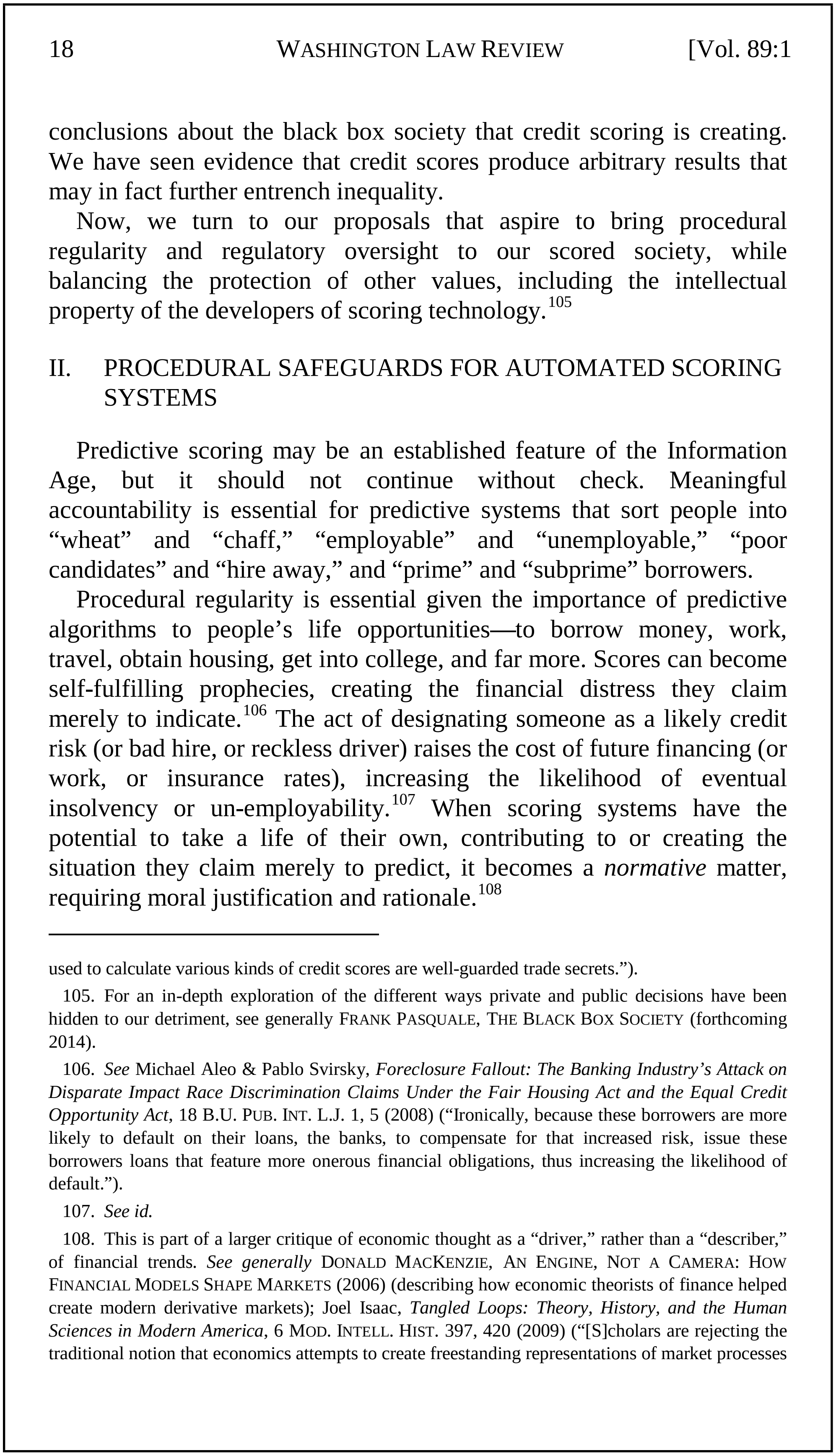
[100] *Id.* § 1681g(f)(C).

[101] *Id.*

[102] *Cf.* Philip Morris v. Reilly, 312 F.3d 24, 47 (1st Cir. 2002) (holding that the state could require revelation of ingredients, but not how much of each was in the cigarettes). The tobacco company in *Reilly* successfully raised a constitutional challenge, alleging the "taking" of a trade secret. *Id.*

[103] Credit histories appear on "consumer reports," as defined by the FCRA. *See* 15 U.S.C. § 1681a(d); Plaintiff's Rule 26(a)(2) Expert Witness Report, Ellis v. Grant & Weber, 2006 WL 3338624 (C.D. Cal. July 26, 2005) (No. CV-04-2007-CAS).

[104] *See* Hendricks, *supra* note 23, at 34 ("Like the recipe for Coca-Cola, the precise formulas



conclusions about the black box society that credit scoring is creating. We have seen evidence that credit scores produce arbitrary results that may in fact further entrench inequality.

Now, we turn to our proposals that aspire to bring procedural regularity and regulatory oversight to our scored society, while balancing the protection of other values, including the intellectual property of the developers of scoring technology.[105]

II. PROCEDURAL SAFEGUARDS FOR AUTOMATED SCORING SYSTEMS

Predictive scoring may be an established feature of the Information Age, but it should not continue without check. Meaningful accountability is essential for predictive systems that sort people into “wheat” and “chaff,” “employable” and “unemployable,” “poor candidates” and “hire away,” and “prime” and “subprime” borrowers.

Procedural regularity is essential given the importance of predictive algorithms to people’s life opportunities—to borrow money, work, travel, obtain housing, get into college, and far more. Scores can become self-fulfilling prophecies, creating the financial distress they claim merely to indicate.[106] The act of designating someone as a likely credit risk (or bad hire, or reckless driver) raises the cost of future financing (or work, or insurance rates), increasing the likelihood of eventual insolvency or un-employability.[107] When scoring systems have the potential to take a life of their own, contributing to or creating the situation they claim merely to predict, it becomes a *normative* matter, requiring moral justification and rationale.[108]

used to calculate various kinds of credit scores are well-guarded trade secrets.”).

[105] For an in-depth exploration of the different ways private and public decisions have been hidden to our detriment, see generally FRANK PASQUALE, THE BLACK BOX SOCIETY (forthcoming 2014).

[106] *See* Michael Aleo & Pablo Svirsky, *Foreclosure Fallout: The Banking Industry’s Attack on Disparate Impact Race Discrimination Claims Under the Fair Housing Act and the Equal Credit Opportunity Act*, 18 B.U. PUB. INT. L.J. 1, 5 (2008) (“Ironically, because these borrowers are more likely to default on their loans, the banks, to compensate for that increased risk, issue these borrowers loans that feature more onerous financial obligations, thus increasing the likelihood of default.”).

[107] *See id.*

[108] This is part of a larger critique of economic thought as a “driver,” rather than a “describer,” of financial trends. *See generally* DONALD MACKENZIE, AN ENGINE, NOT A CAMERA: HOW FINANCIAL MODELS SHAPE MARKETS (2006) (describing how economic theorists of finance helped create modern derivative markets); Joel Isaac, *Tangled Loops: Theory, History, and the Human Sciences in Modern America*, 6 MOD. INTELL. HIST. 397, 420 (2009) (“[S]cholars are rejecting the traditional notion that economics attempts to create freestanding representations of market processes

Scoring systems should be subject to fairness requirements that reflect their centrality in people's lives. Private scoring systems should be as understandable to regulators as to firms' engineers. However well an "invisible hand" coordinates economic activity generally speaking, markets depend on reliable information about the practices of firms that finance, rank, and rate consumers. Brandishing quasi-governmental authority to determine which individuals are worthy of financial backing, private scoring systems need to be held to a higher standard than the average firm.

One of the great accomplishments of the legal order was holding the sovereign accountable for decisionmaking and giving subjects basic rights, in breakthroughs stretching from Runnymede to the Glorious Revolution of 1688 to the American Revolution. New algorithmic decisionmakers are sovereign over important aspects of individual lives. If law and due process are absent from this field, we are essentially paving the way to a new feudal order of unaccountable reputational intermediaries.[109]

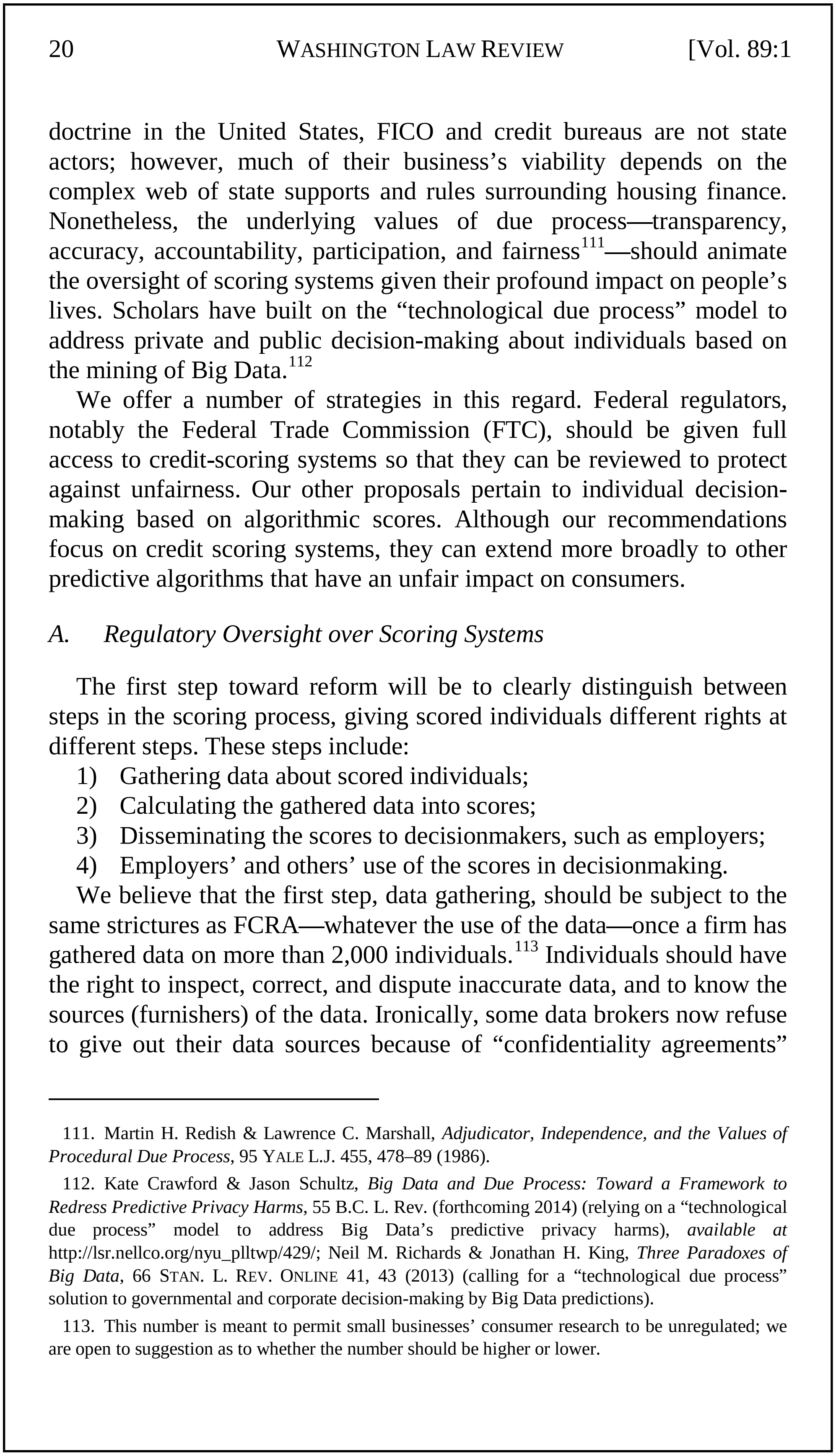
How should we accomplish accountability? Protections could draw insights from what one of us has called "technological due process"—procedures ensuring that predictive algorithms live up to some standard of review and revision to ensure their fairness and accuracy.[110] Procedural protections should apply not only to the scoring algorithms themselves (a kind of technology-driven rulemaking), but also to individual decisions based on algorithmic predictions (technology-driven adjudication).

This is not to suggest that full due process guarantees are required as a matter of current law. Given the etiolated state of "state action"

(which economic sociologists must then insist leaves out power, or cultural context, or the fullness of human agency)."). Some commentators have argued that we need to "recognize economics not as a (misguided) science of capitalism but as its technology, that is, as one of the active ingredients in the production and reproduction of the market order." Marion Fourcade, *Theories of Markets and Theories of Society*, 50 AM. BEHAV. SCI. 1015, 1025 (2007).

[109] Our proposal for basic rights of citizens vis-á-vis scoring systems also finds support in the work of other scholars concerned about the extraordinary power of private companies. *See, e.g.*, LORI ANDREWS, I KNOW WHO YOU ARE AND I SAW WHAT YOU DID: SOCIAL NETWORKS AND THE DEATH OF PRIVACY 189–91 (2012) (concluding with a proposal for a "Social Network Constitution"); REBECCA MACKINNON, CONSENT OF THE NETWORKED 240–41 (2012) (proposing ten principles of network governance); Jeffrey Rosen, *Madison's Privacy Blind Spot*, N.Y. TIMES, Jan. 19, 2014 (Sunday), at 5 ("What Americans may now need is a constitutional amendment to prohibit unreasonable searches and seizures of our persons and electronic effects, whether by the government or by private corporations like Google and AT&T. . . . [O]ur rights to enjoy liberty, and to obtain happiness and safety at the same time, are threatened as much by corporate as government surveillance.").

[110] *See generally* Citron, *supra* note 20.



doctrine in the United States, FICO and credit bureaus are not state actors; however, much of their business's viability depends on the complex web of state supports and rules surrounding housing finance. Nonetheless, the underlying values of due process—transparency, accuracy, accountability, participation, and fairness[111]—should animate the oversight of scoring systems given their profound impact on people's lives. Scholars have built on the "technological due process" model to address private and public decision-making about individuals based on the mining of Big Data.[112]

We offer a number of strategies in this regard. Federal regulators, notably the Federal Trade Commission (FTC), should be given full access to credit-scoring systems so that they can be reviewed to protect against unfairness. Our other proposals pertain to individual decision-making based on algorithmic scores. Although our recommendations focus on credit scoring systems, they can extend more broadly to other predictive algorithms that have an unfair impact on consumers.

A. Regulatory Oversight over Scoring Systems

The first step toward reform will be to clearly distinguish between steps in the scoring process, giving scored individuals different rights at different steps. These steps include:

1) Gathering data about scored individuals;
2) Calculating the gathered data into scores;
3) Disseminating the scores to decisionmakers, such as employers;
4) Employers' and others' use of the scores in decisionmaking.

We believe that the first step, data gathering, should be subject to the same strictures as FCRA—whatever the use of the data—once a firm has gathered data on more than 2,000 individuals.[113] Individuals should have the right to inspect, correct, and dispute inaccurate data, and to know the sources (furnishers) of the data. Ironically, some data brokers now refuse to give out their data sources because of "confidentiality agreements"

[111] Martin H. Redish & Lawrence C. Marshall, *Adjudicator, Independence, and the Values of Procedural Due Process*, 95 YALE L.J. 455, 478–89 (1986).

[112] Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. Rev. (forthcoming 2014) (relying on a "technological due process" model to address Big Data's predictive privacy harms), *available at* http://lsr.nellco.org/nyu_plltwp/429/; Neil M. Richards & Jonathan H. King, *Three Paradoxes of Big Data*, 66 STAN. L. REV. ONLINE 41, 43 (2013) (calling for a "technological due process" solution to governmental and corporate decision-making by Big Data predictions).

[113] This number is meant to permit small businesses' consumer research to be unregulated; we are open to suggestion as to whether the number should be higher or lower.

with sources.[114] That position (hiding behind privacy interests to violate consumer privacy) would not stand for consumer reporting agencies covered by FCRA. It should not stand for data brokers and the like.

Second, at the calculation of data stage, ideally such calculations would be public, and all processes (whether driven by AI or other computing) would be inspectable. In some cases, the trade secrets may merit protection, and only a dedicated, closed review should be available. But in general, we need to switch the default in situations like this away from an assumption of secrecy, and toward the expectation that people deserve to know how they are rated and ranked.

The third stage is more difficult, as it begins to implicate First Amendment issues. Given the Supreme Court's ruling in *Sorrell v. IMS Health Inc.*[115] and other rulings in cases involving the regulation of ranking systems,[116] courts may look askance at rules that limit the dissemination of data or scores.[117] Nevertheless, scored individuals should be notified when scores or data are communicated to an entity. That notification only *increases* speech; it does not restrict or censor communication. Coerced speech can implicate the First Amendment, but like Professor Neil Richards, we do not understand *Sorrell* to lay down a blanket rule that all data is speech.[118] Transparency requirements are consistent with First Amendment doctrine.

The fourth and final stage is the most controversial. We believe that—given the sensitivity of scoring and their disparate impact on vulnerable populations—scoring systems should be subject to licensing and audit requirements when they enter critical settings like employment,

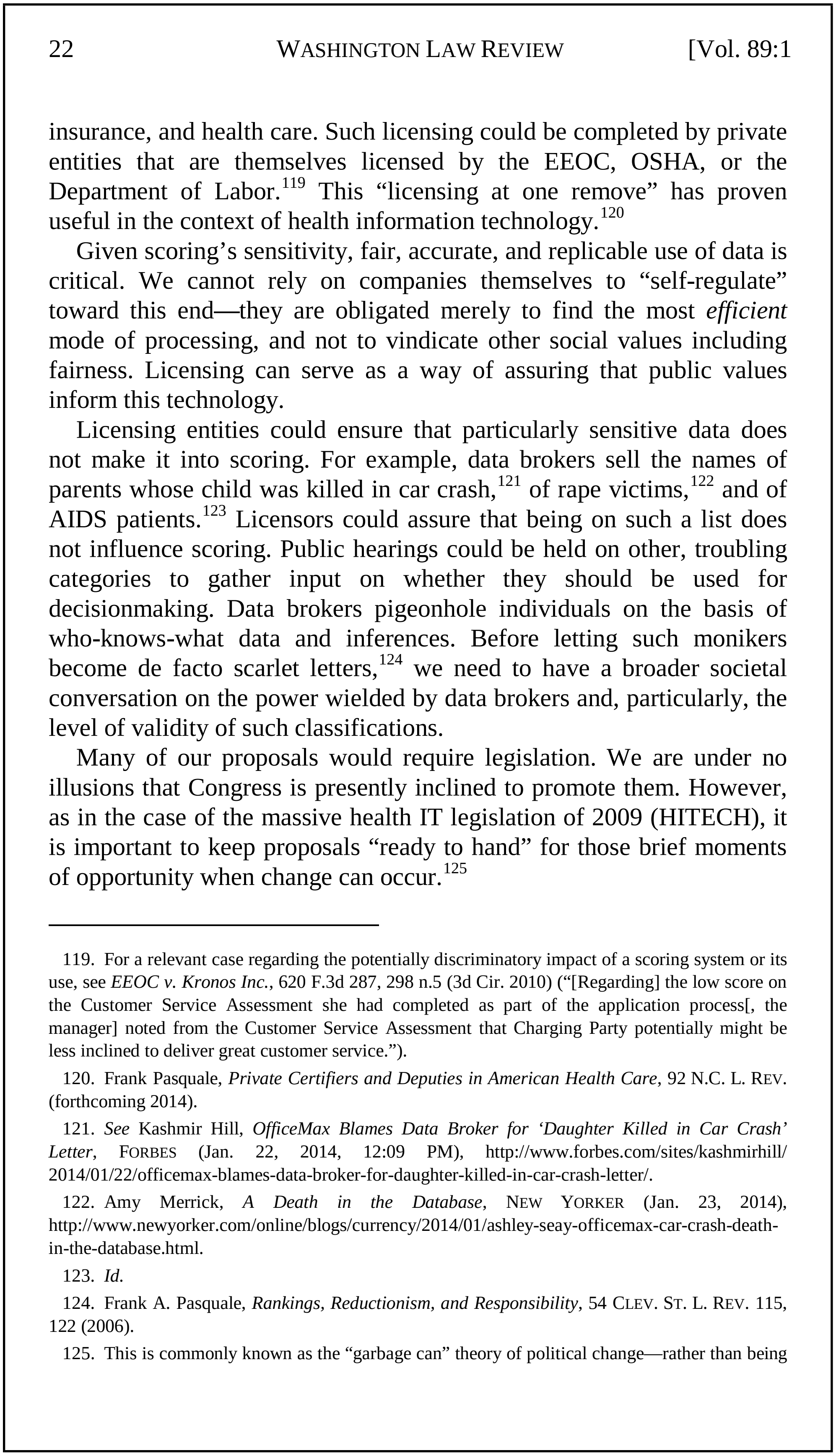
[114] Casey Johnston, *Data Brokers Won't Even Tell the Government How It Uses, Sells Your Data*, ARS TECHNICA (Dec. 21, 2013, 12:07 PM), http://arstechnica.com/business/2013/12/data-brokers-wont-even-tell-the-government-how-it-uses-sells-your-data/.

[115] __U.S.__, 131 S. Ct. 2653 (2011).

[116] *See, e.g.*, Pasquale, *Beyond Innovation and Competition*, *supra* note 34, at 117–19 (discussing the successful First Amendment defense of the Avvo lawyer ratings site).

[117] *Sorrel*, 131 S. Ct. at 2670–72 (holding that drug companies have a constitutional right to access certain types of data without undue state interference); *see also* NEIL M. RICHARDS, INTELLECTUAL PRIVACY: CIVIL LIBERTIES AND INFORMATION IN A DIGITAL AGE ch. 5 (forthcoming 2014) (exploring why *Sorrell* does not lay down a blanket rule that all data is speech for purposes of the First Amendment and more narrowly rested on concerns about viewpoint discrimination among other reasons). For a critical description of the stakes of *Sorrell*, see David Orentlicher, *Prescription Data Mining and the Protection of Patients' Interests*, 38 J.L. MED. & ETHICS 74, 81 (2010) ("When people develop relationships with their physicians and pharmacists, they are entitled to the assurance that information about their medical condition will be used for their benefit and not to place their health at risk or to increase their health care costs."); Frank Pasquale, *Grand Bargains for Big Data*, 72 MD. L. REV. 682, 740 (2013); Andrew Tutt, *Software Speech*, 65 STAN. L. REV. ONLINE 73, 75 (2012).

[118] *See* RICHARDS, *supra* note 117, at ch. 5.



insurance, and health care. Such licensing could be completed by private entities that are themselves licensed by the EEOC, OSHA, or the Department of Labor.[119] This "licensing at one remove" has proven useful in the context of health information technology.[120]

Given scoring's sensitivity, fair, accurate, and replicable use of data is critical. We cannot rely on companies themselves to "self-regulate" toward this end—they are obligated merely to find the most *efficient* mode of processing, and not to vindicate other social values including fairness. Licensing can serve as a way of assuring that public values inform this technology.

Licensing entities could ensure that particularly sensitive data does not make it into scoring. For example, data brokers sell the names of parents whose child was killed in car crash,[121] of rape victims,[122] and of AIDS patients.[123] Licensors could assure that being on such a list does not influence scoring. Public hearings could be held on other, troubling categories to gather input on whether they should be used for decisionmaking. Data brokers pigeonhole individuals on the basis of who-knows-what data and inferences. Before letting such monikers become de facto scarlet letters,[124] we need to have a broader societal conversation on the power wielded by data brokers and, particularly, the level of validity of such classifications.

Many of our proposals would require legislation. We are under no illusions that Congress is presently inclined to promote them. However, as in the case of the massive health IT legislation of 2009 (HITECH), it is important to keep proposals "ready to hand" for those brief moments of opportunity when change can occur.[125]

[119] For a relevant case regarding the potentially discriminatory impact of a scoring system or its use, see *EEOC v. Kronos Inc.*, 620 F.3d 287, 298 n.5 (3d Cir. 2010) ("[Regarding] the low score on the Customer Service Assessment she had completed as part of the application process[, the manager] noted from the Customer Service Assessment that Charging Party potentially might be less inclined to deliver great customer service.").

[120] Frank Pasquale, *Private Certifiers and Deputies in American Health Care*, 92 N.C. L. REV. (forthcoming 2014).

[121] *See* Kashmir Hill, *OfficeMax Blames Data Broker for 'Daughter Killed in Car Crash' Letter*, FORBES (Jan. 22, 2014, 12:09 PM), http://www.forbes.com/sites/kashmirhill/2014/01/22/officemax-blames-data-broker-for-daughter-killed-in-car-crash-letter/.

[122] Amy Merrick, *A Death in the Database*, NEW YORKER (Jan. 23, 2014), http://www.newyorker.com/online/blogs/currency/2014/01/ashley-seay-officemax-car-crash-death-in-the-database.html.

[123] *Id.*

[124] Frank A. Pasquale, *Rankings, Reductionism, and Responsibility*, 54 CLEV. ST. L. REV. 115, 122 (2006).

[125] This is commonly known as the "garbage can" theory of political change—rather than being

Fortunately, the Federal Trade Commission does have statutory authority to move forward on several parts of the “scored society” agenda. The FTC can oversee credit-scoring systems under its authority to combat “unfair” trade practices under Section 5 of the Federal Trade Commission Act.[126] It can use this authority to develop much more robust oversight over credit scoring, which could then be a model for legislation for other scoring entities (or for state consumer protection authorities and state attorneys general with authority to promote fair information practices).

“Unfair” commercial practices involve conduct that substantially harms consumers, or threatens to substantially harm consumers, which consumers cannot reasonably avoid, and where the harm outweighs the benefits.[127] In 2008, the FTC invoked its unfairness authority against a credit provider for basing credit reductions on an undisclosed behavioral scoring model that penalized consumers for using their credit cards for certain transactions, such as personal counseling.[128]

The FTC’s concerns about predictive algorithms have escalated with their increasing use. In March 2014, the FTC is hosting a panel of experts to discuss the private sector’s use of algorithmic scores to make decisions about individuals, including individuals’ credit risk with certain transactions, likelihood to take medication, and influence over others based on networked activities.[129] The FTC has identified the following topics for discussion:

- How are companies utilizing these predictive scores?
- How accurate are these scores and the underlying data

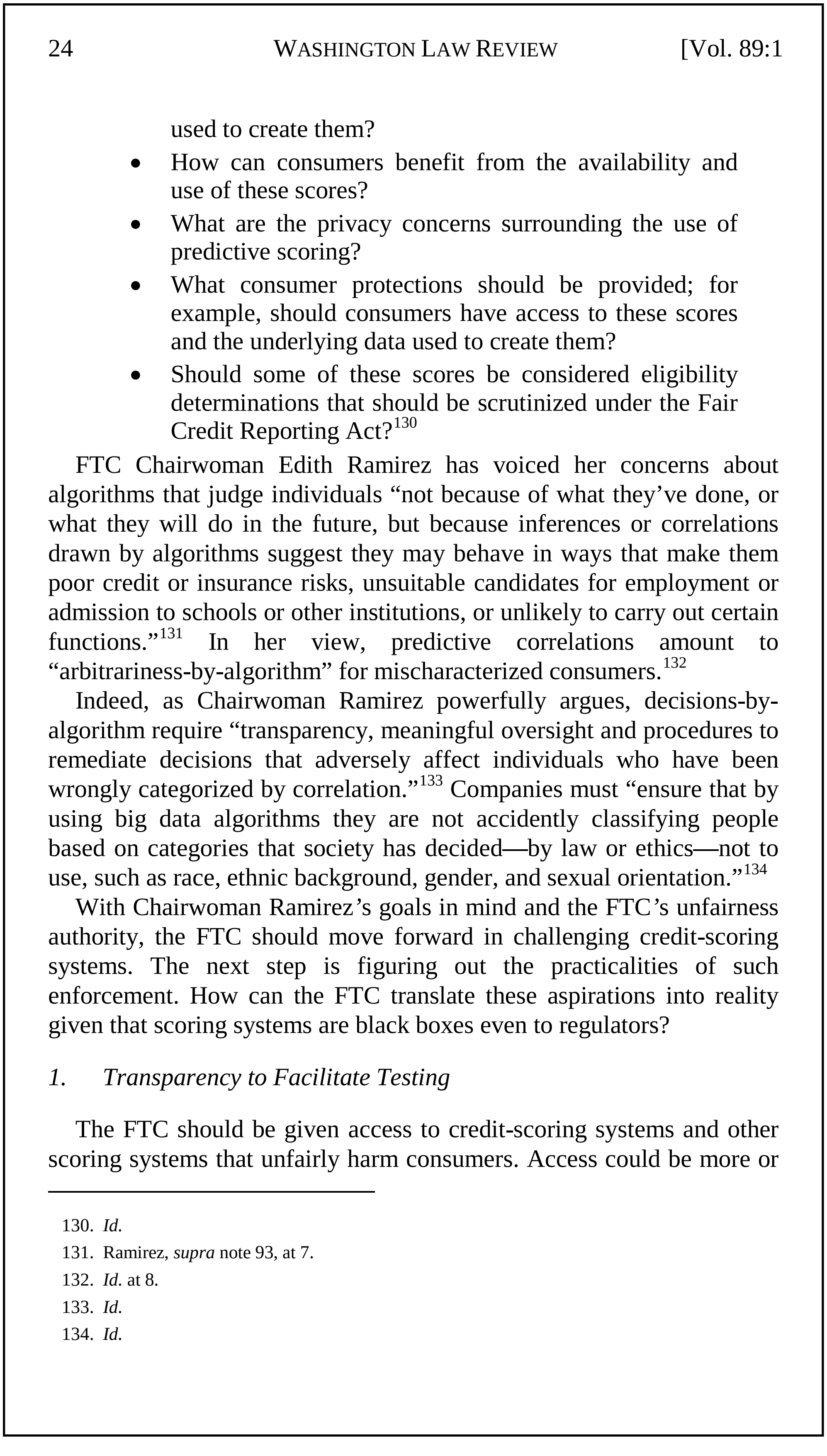
rationally planned, most legislative efforts depend on whatever plans are at hand. J. Bendor et al., *Recycling the Garbage Can: An Assessment of the Research Program*, 95 AM. POL. SCI. REV. 95, 169 (2001).

[126] *See* Federal Trade Commission Act § 5, 15 U.S.C. § 45 (2012). *See generally A Brief Overview of the Federal Trade Commission’s Investigative and Law Enforcement Authority*, FED. TRADE COMM’N, http://www.ftc.gov/about-ftc/what-we-do/enforcement-authority (last updated July 2008).

[127] 15 U.S.C. § 45(n) (2012).

[128] Stipulated Order for Permanent Injunction and Other Equitable Relief Against Defendant CompuCredit Corp., FTC v. CompuCredit Corp., No. 1:08-CV-1976-BBM-RGV (N.D. Ga. Dec. 19, 2008), *available at* http://www.ftc.gov/sites/default/files/documents/cases/2008/12/081219compucreditstiporder.pdf. For a compelling account of the crucial role that the FTC plays in regulating unfair consumer practices and establishing a common law of privacy, see Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. (forthcoming 2014), *available at* http://ssrn.com/abstract=2312913 (last updated Oct. 29, 2013).

[129] *See Spring Privacy Series: Alternative Scoring Products*, FED. TRADE COMM’N, http://www.ftc.gov/news-events/events-calendar/2014/03/spring-privacy-series-alternative-scoring-products (last visited Feb. 11, 2014).



used to create them?

- How can consumers benefit from the availability and use of these scores?
- What are the privacy concerns surrounding the use of predictive scoring?
- What consumer protections should be provided; for example, should consumers have access to these scores and the underlying data used to create them?
- Should some of these scores be considered eligibility determinations that should be scrutinized under the Fair Credit Reporting Act?[130]

FTC Chairwoman Edith Ramirez has voiced her concerns about algorithms that judge individuals “not because of what they’ve done, or what they will do in the future, but because inferences or correlations drawn by algorithms suggest they may behave in ways that make them poor credit or insurance risks, unsuitable candidates for employment or admission to schools or other institutions, or unlikely to carry out certain functions.”[131] In her view, predictive correlations amount to “arbitrariness-by-algorithm” for mischaracterized consumers.[132]

Indeed, as Chairwoman Ramirez powerfully argues, decisions-by-algorithm require “transparency, meaningful oversight and procedures to remediate decisions that adversely affect individuals who have been wrongly categorized by correlation.”[133] Companies must “ensure that by using big data algorithms they are not accidently classifying people based on categories that society has decided—by law or ethics—not to use, such as race, ethnic background, gender, and sexual orientation.”[134]

With Chairwoman Ramirez’s goals in mind and the FTC’s unfairness authority, the FTC should move forward in challenging credit-scoring systems. The next step is figuring out the practicalities of such enforcement. How can the FTC translate these aspirations into reality given that scoring systems are black boxes even to regulators?

1. *Transparency to Facilitate Testing*

The FTC should be given access to credit-scoring systems and other scoring systems that unfairly harm consumers. Access could be more or

[130] *Id.*

[131] Ramirez, *supra* note 93, at 7.

[132] *Id.* at 8.

[133] *Id.*

[134] *Id.*

less episodic depending on the extent of unfairness exhibited by the scoring system. Biannual audits would make sense for most scoring systems; more frequent monitoring would be necessary for those which had engaged in troubling conduct.[135]

We should be particularly focused on scoring systems which rank and rate individuals who can do little or nothing to protect themselves. The FTC's expert technologists[136] could test scoring systems for bias, arbitrariness, and unfair mischaracterizations. To do so, they would need to view not only the datasets mined by scoring systems[137] but also the source code and programmers' notes describing the variables, correlations, and inferences embedded in the scoring systems' algorithms.[138]

For the review to be meaningful in an era of great technological change, the FTC's technical experts must be able to meaningfully assess systems whose predictions change pursuant to AI logic. They should permitted to test systems to detect patterns and correlations tied to classifications that are already suspect under American law, such as race, nationality, sexual orientation, and gender. Scoring systems should be run through testing suites that run expected and unexpected hypothetical scenarios designed by policy experts.[139] Testing reflects the norm of proper software development, and would help detect both programmers' potential bias and bias emerging from the AI system's evolution.[140]

2. *Risk Assessment Reports and Recommendations*

Once the FTC evaluates credit-scoring systems to detect

[135] *See* Helen Nissenbaum, *Accountability in a Computerized Society*, 2 SCI. & ENGINEERING ETHICS 25, 37 (1996) (describing commentators' calls for "simpler design, a modular approach to system building, meaningful quality assurance, independent auditing, built-in redundancy, and excellent documentation").

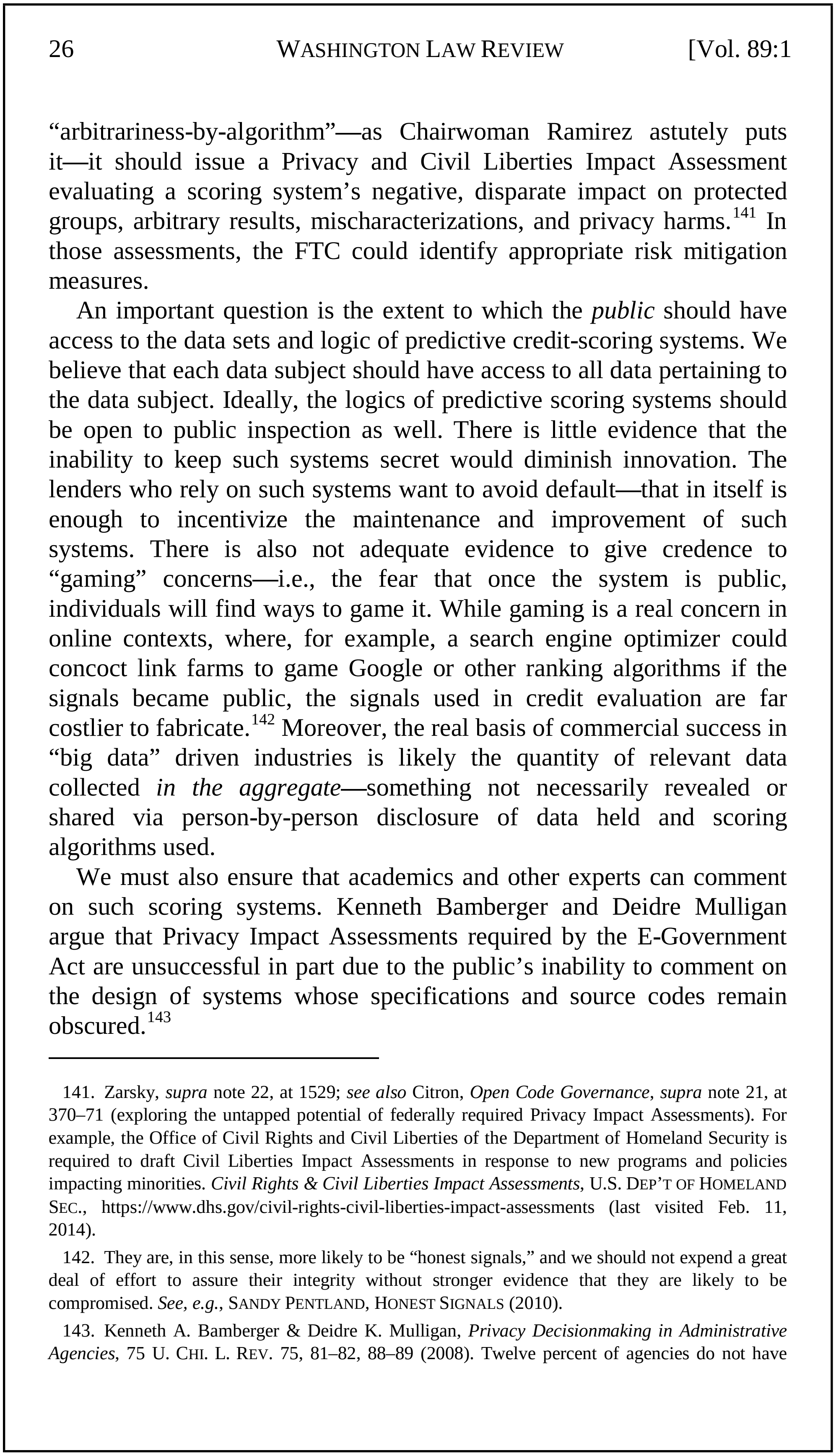
[136] The FTC's Senior Technologist position has been filled by esteemed computer scientists Professor Edward Felten of Princeton University, Professor Steven Bellovin of Columbia University, and now by Professor LaTanya Sweeney of Harvard University.

[137] *See, e.g.*, Zarsky, *supra* note 22, at 1520.

[138] We thank Ed Felten for suggesting that oversight of automated systems include access to programmers' notes for the purpose of assessing source code. Ed Felten, Comment to Danielle Citron, Technological Due Process Lecture at Princeton University Center on Information Technology Policy Lecture Series (Apr. 30, 2009); *see also Danielle Citron: Technological Due Process*, CTR. FOR INFO. TECH. POL'Y, https://citp.princeton.edu/event/citron/ (last visited Feb. 11, 2014). The question we shall soon address is whether the public generally and affected individuals specifically should *also* have access to the data sets and logic behind predictive algorithms.

[139] Citron, *supra* note 20, at 1310.

[140] Batya Friedman & Helen Nissenbaum, *Bias in Computer Systems*, 14 ACM TRANSACTIONS ON INFO. SYSTEMS 330, 334 (1996).



"arbitrariness-by-algorithm"—as Chairwoman Ramirez astutely puts it—it should issue a Privacy and Civil Liberties Impact Assessment evaluating a scoring system's negative, disparate impact on protected groups, arbitrary results, mischaracterizations, and privacy harms.[141] In those assessments, the FTC could identify appropriate risk mitigation measures.

An important question is the extent to which the *public* should have access to the data sets and logic of predictive credit-scoring systems. We believe that each data subject should have access to all data pertaining to the data subject. Ideally, the logics of predictive scoring systems should be open to public inspection as well. There is little evidence that the inability to keep such systems secret would diminish innovation. The lenders who rely on such systems want to avoid default—that in itself is enough to incentivize the maintenance and improvement of such systems. There is also not adequate evidence to give credence to "gaming" concerns—i.e., the fear that once the system is public, individuals will find ways to game it. While gaming is a real concern in online contexts, where, for example, a search engine optimizer could concoct link farms to game Google or other ranking algorithms if the signals became public, the signals used in credit evaluation are far costlier to fabricate.[142] Moreover, the real basis of commercial success in "big data" driven industries is likely the quantity of relevant data collected *in the aggregate*—something not necessarily revealed or shared via person-by-person disclosure of data held and scoring algorithms used.

We must also ensure that academics and other experts can comment on such scoring systems. Kenneth Bamberger and Deidre Mulligan argue that Privacy Impact Assessments required by the E-Government Act are unsuccessful in part due to the public's inability to comment on the design of systems whose specifications and source codes remain obscured.[143]

[141] Zarsky, *supra* note 22, at 1529; *see also* Citron, *Open Code Governance*, *supra* note 21, at 370–71 (exploring the untapped potential of federally required Privacy Impact Assessments). For example, the Office of Civil Rights and Civil Liberties of the Department of Homeland Security is required to draft Civil Liberties Impact Assessments in response to new programs and policies impacting minorities. *Civil Rights & Civil Liberties Impact Assessments*, U.S. DEP'T OF HOMELAND SEC., https://www.dhs.gov/civil-rights-civil-liberties-impact-assessments (last visited Feb. 11, 2014).

[142] They are, in this sense, more likely to be "honest signals," and we should not expend a great deal of effort to assure their integrity without stronger evidence that they are likely to be compromised. *See, e.g.*, SANDY PENTLAND, HONEST SIGNALS (2010).

[143] Kenneth A. Bamberger & Deidre K. Mulligan, *Privacy Decisionmaking in Administrative Agencies*, 75 U. CHI. L. REV. 75, 81–82, 88–89 (2008). Twelve percent of agencies do not have

As Tal Zarsky argues, the public could be informed about the datasets that predictive systems mine without generating significant social risks.[144] Zarsky demonstrates that—when it comes to "the collection of data and aggregation of datasets"—it is evident that "providing information regarding the kinds and forms of data and databases used in the analysis . . . generate[s] limited social risks . . . [usually only in the context of] secretive governmental datasets."[145]

The more difficult question concerns whether scoring systems' source code, algorithmic predictions, and modeling should be transparent to affected individuals and ultimately the public at large. Neil Richards and Jonathan King astutely explain that "there are legitimate arguments for some level of big data secrecy," including concerns "connected to highly sensitive intellectual property and national security assets."[146] But these concerns are more than outweighed by the threats to human dignity posed by pervasive, secret, and automated scoring systems. At the very least, individuals should have a meaningful form of notice and a chance to challenge predictive scores that harm their ability to obtain credit, jobs, housing, and other important opportunities.

B. Protections for Individuals

In constructing strategies for technological due process in scoring contexts, it is helpful to consider the sort of notice individuals are owed when governmental systems make adverse decisions about them. Under the Due Process Clause, notice must be "reasonably calculated" to inform individuals of the government's claims against them.[147] The sufficiency of notice depends upon its ability to inform affected individuals about the issues to be decided, the evidence supporting the government's position, and the agency's decisional process.[148] Clear notice decreases the likelihood that agency action will rest upon "incorrect or misleading factual premises or on the misapplication of rules."[149]

written processes or policies for all listed aspects of Privacy Impact Assessment (PIA) and sixteen percent of systems covered by the PIA requirement did not have a complete or current PIA. *Id.* at 81.

[144] Zarsky, *supra* note 22, at 1524 (exploring the practical and normative implications of varying kinds of transparency for governmental predictive systems).

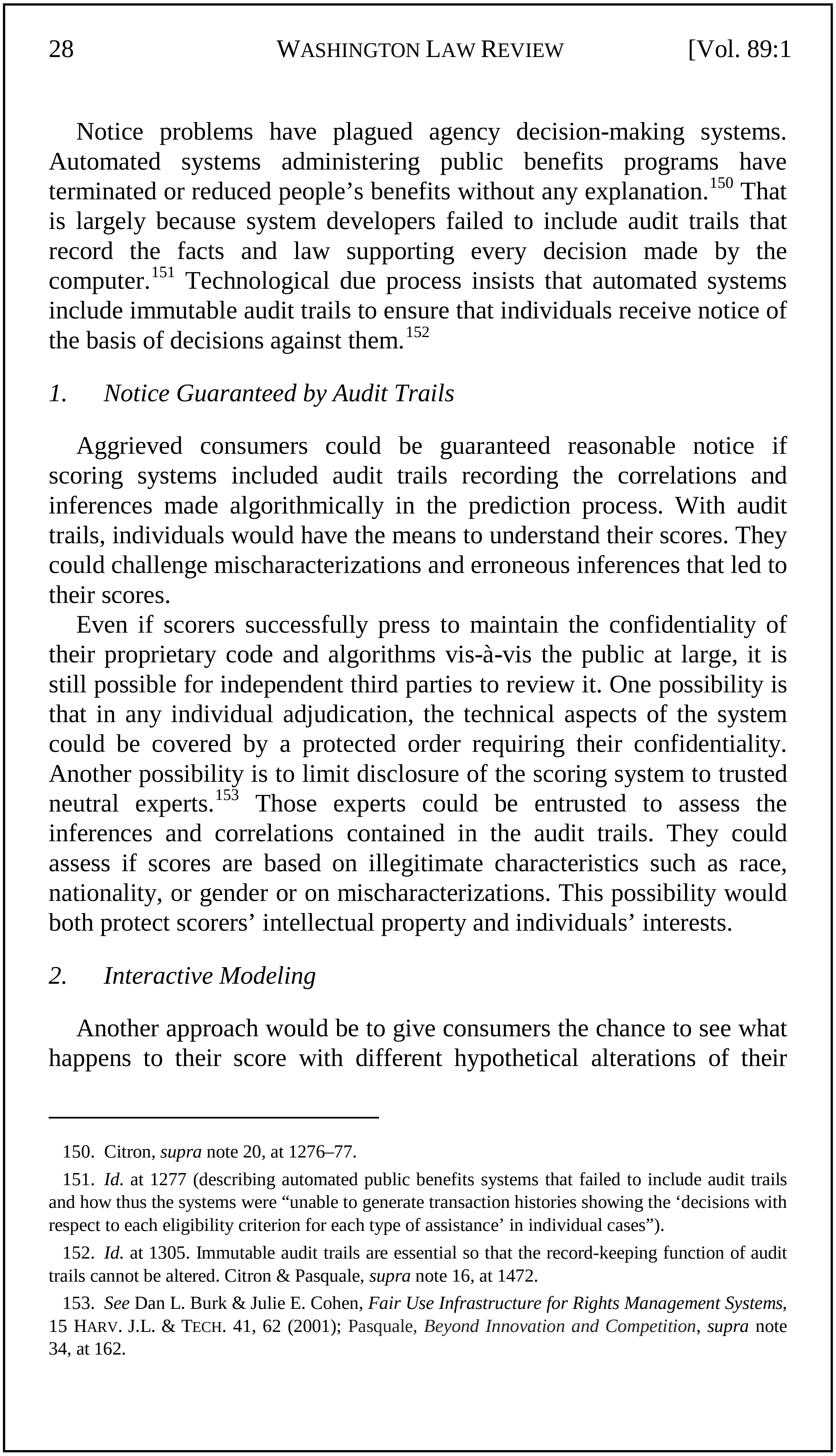
[145] *Id.*

[146] Richards & King, *supra* note 112, at 43.

[147] Dusenbery v. United States, 534 U.S. 161, 168 (2002).

[148] JERRY L. MASHAW, DUE PROCESS IN THE ADMINISTRATIVE STATE 176 (1985).

[149] Goldberg v. Kelly, 397 U.S. 254, 268 (1970).



Notice problems have plagued agency decision-making systems. Automated systems administering public benefits programs have terminated or reduced people's benefits without any explanation.[150] That is largely because system developers failed to include audit trails that record the facts and law supporting every decision made by the computer.[151] Technological due process insists that automated systems include immutable audit trails to ensure that individuals receive notice of the basis of decisions against them.[152]

1. *Notice Guaranteed by Audit Trails*

Aggrieved consumers could be guaranteed reasonable notice if scoring systems included audit trails recording the correlations and inferences made algorithmically in the prediction process. With audit trails, individuals would have the means to understand their scores. They could challenge mischaracterizations and erroneous inferences that led to their scores.

Even if scorers successfully press to maintain the confidentiality of their proprietary code and algorithms vis-à-vis the public at large, it is still possible for independent third parties to review it. One possibility is that in any individual adjudication, the technical aspects of the system could be covered by a protected order requiring their confidentiality. Another possibility is to limit disclosure of the scoring system to trusted neutral experts.[153] Those experts could be entrusted to assess the inferences and correlations contained in the audit trails. They could assess if scores are based on illegitimate characteristics such as race, nationality, or gender or on mischaracterizations. This possibility would both protect scorers' intellectual property and individuals' interests.

2. *Interactive Modeling*

Another approach would be to give consumers the chance to see what happens to their score with different hypothetical alterations of their

[150] Citron, *supra* note 20, at 1276–77.

[151] *Id.* at 1277 (describing automated public benefits systems that failed to include audit trails and how thus the systems were "unable to generate transaction histories showing the 'decisions with respect to each eligibility criterion for each type of assistance' in individual cases").

[152] *Id.* at 1305. Immutable audit trails are essential so that the record-keeping function of audit trails cannot be altered. Citron & Pasquale, *supra* note 16, at 1472.

[153] *See* Dan L. Burk & Julie E. Cohen, *Fair Use Infrastructure for Rights Management Systems*, 15 HARV. J.L. & TECH. 41, 62 (2001); Pasquale, *Beyond Innovation and Competition*, *supra* note 34, at 162.

credit histories. Imagine an interface where each aspect of a person's credit history is represented on a wiki.[154] To make it more concrete, picture a consumer who is facing a dilemma. She sees on her credit report that she has a bill that is thirty days overdue. She could secure a payday loan to pay the bill, but she'd face a usurious interest rate if she takes that option. She can probably earn enough money working overtime to pay the bill herself in forty days. Software could give her an idea of the relative merits of either course. If her score dropped by 100 points when a bill went unpaid for a total of sixty days, she would be much more likely to opt for the payday loan than if a mere five points were deducted for that term of delinquency.

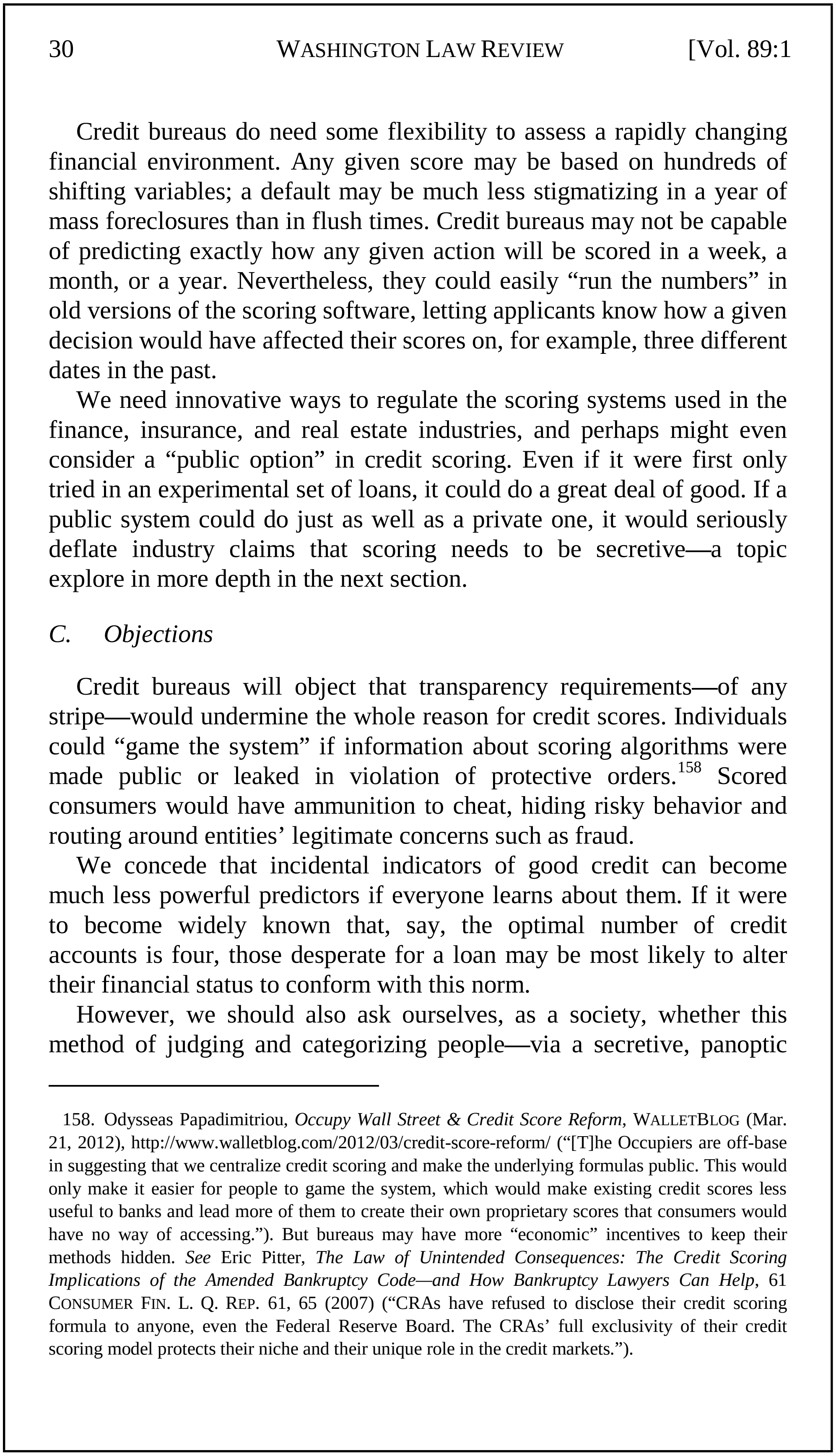
Just as the authors of the children's series *Choose Your Own Adventure* helped pave the way to the cornucopia of interactive entertainment now offered today,[155] so, too, might creative customer relations demystify credit scoring. Interactive modeling, known as "feedback and control," has been successfully deployed in other technical contexts by a "values in design" movement.[156] It has promoted automated systems that give individuals more of a sense of how future decisions will affect their evaluation. For example, Canada's Immigration Bureau lets individuals enter various scenarios into a preliminary "test" for qualification as a permanent resident.[157] The digital interface allows users to estimate how different decisions will affect their potential to become a Canadian citizen. Learning French or earning a graduate degree can be a great help to those in their thirties; on the other hand, some over sixty years old can do "everything right" and still end up with too few points to apply successfully. The public scorecard does not guarantee anyone admittance, and is revised over time. Nevertheless, it provides a rough outline of what matters to the scoring process, and how much.

[154] For general information on wikis, see Daniel Nations, *What is a Wiki?*, ABOUT.COM, http://webtrends.about.com/od/wiki/a/what_is_a_wiki.htm (last visited Feb. 11, 2014).

[155] Grady Hendrix, *Choose Your Own Adventure*, SLATE (Feb. 18, 2011, 7:08 AM), http://www.slate.com/id/2282786/.

[156] Comments of Deirdre K. Mulligan, Professor, Univ. of Calif. at Berkeley & Nicholas P. Doty in Response to the National Telecommunications & Information Administration's Request for Comments on the Multistakeholder Process To Develop Consumer Data Privacy Codes of Conduct, Docket No. 120214135-2135-01, at 11 (May 18, 2012), *available at* http://www.ntia.doc.gov/files/ntia/mulligan_doty_comments.pdf. *See generally* HELEN NISSENBAUM, PRIVACY IN CONTEXT: TECHNOLOGY, POLICY, AND THE INTEGRITY OF SOCIAL LIFE (2010); PROFILING THE EUROPEAN CITIZEN 67 (Mireille Hildebrandt & Serge Gutwirth eds., 2008).

[157] *Determine Your Eligibility — Federal Skilled Workers*, GOV'T OF CANADA, http://www.cic.gc.ca/english/immigrate/skilled/apply-who.asp (last updated June 20, 2013).



Credit bureaus do need some flexibility to assess a rapidly changing financial environment. Any given score may be based on hundreds of shifting variables; a default may be much less stigmatizing in a year of mass foreclosures than in flush times. Credit bureaus may not be capable of predicting exactly how any given action will be scored in a week, a month, or a year. Nevertheless, they could easily "run the numbers" in old versions of the scoring software, letting applicants know how a given decision would have affected their scores on, for example, three different dates in the past.

We need innovative ways to regulate the scoring systems used in the finance, insurance, and real estate industries, and perhaps might even consider a "public option" in credit scoring. Even if it were first only tried in an experimental set of loans, it could do a great deal of good. If a public system could do just as well as a private one, it would seriously deflate industry claims that scoring needs to be secretive—a topic explore in more depth in the next section.

C. Objections

Credit bureaus will object that transparency requirements—of any stripe—would undermine the whole reason for credit scores. Individuals could "game the system" if information about scoring algorithms were made public or leaked in violation of protective orders.[158] Scored consumers would have ammunition to cheat, hiding risky behavior and routing around entities' legitimate concerns such as fraud.

We concede that incidental indicators of good credit can become much less powerful predictors if everyone learns about them. If it were to become widely known that, say, the optimal number of credit accounts is four, those desperate for a loan may be most likely to alter their financial status to conform with this norm.

However, we should also ask ourselves, as a society, whether this method of judging and categorizing people—via a secretive, panoptic

[158] Odysseas Papadimitriou, *Occupy Wall Street & Credit Score Reform*, WALLETBLOG (Mar. 21, 2012), http://www.walletblog.com/2012/03/credit-score-reform/ ("[T]he Occupiers are off-base in suggesting that we centralize credit scoring and make the underlying formulas public. This would only make it easier for people to game the system, which would make existing credit scores less useful to banks and lead more of them to create their own proprietary scores that consumers would have no way of accessing."). But bureaus may have more "economic" incentives to keep their methods hidden. *See* Eric Pitter, *The Law of Unintended Consequences: The Credit Scoring Implications of the Amended Bankruptcy Code—and How Bankruptcy Lawyers Can Help*, 61 CONSUMER FIN. L. Q. REP. 61, 65 (2007) ("CRAs have refused to disclose their credit scoring formula to anyone, even the Federal Reserve Board. The CRAs' full exclusivity of their credit scoring model protects their niche and their unique role in the credit markets.").

sort—is appropriate. It has already contributed to one of the greatest financial crises in American history, legitimizing widespread subprime lending by purporting to scientifically rank individuals' creditworthiness with extraordinary precision. Secretive credit scoring can needlessly complicate the social world, lend a patina of objectivity to dangerous investment practices, and encode discriminatory practices in impenetrable algorithms.[159]

The benefits of secrecy are murkier than these costs. Moreover, the secrecy of credit scoring can impede incremental innovation: how can outsiders develop better scoring systems if they have no way of accessing current ones? Secret credit scoring can undermine the public good, since opaque methods of scoring make it difficult for those who feel—and quite possibly are—wronged to press their case.

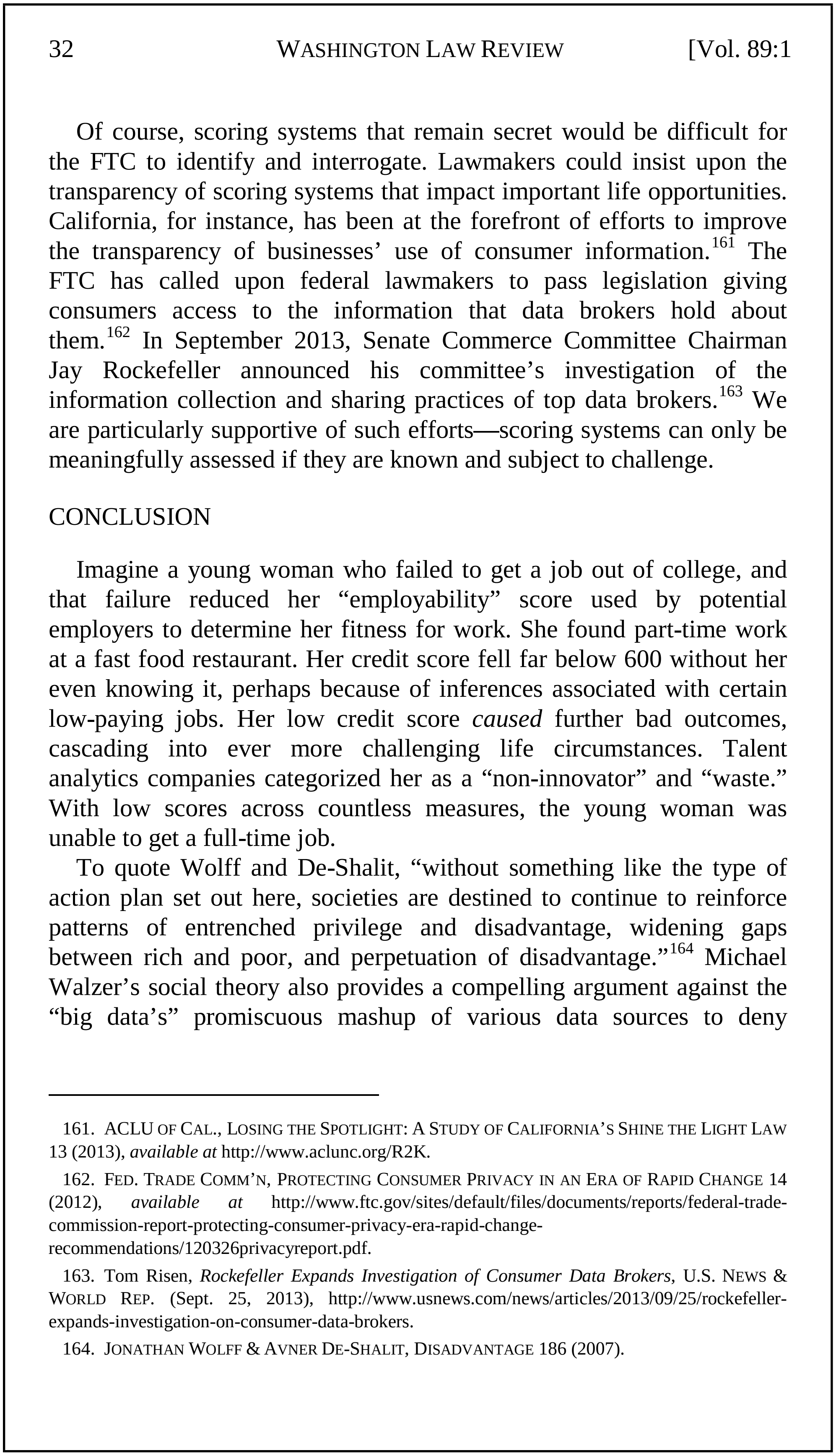
If scorers can produce evidence about the bad effects of publicity, that might justify keeping the correlations, inferences, and logic of scoring algorithms from the public at large. But that logic would not apply to the FTC or third-party experts who would be bound to keep proprietary information confidential.

Another objection is that our proposal only works when the very existence of scoring systems is public knowledge, as in the case of credit scores. In non-credit contexts, entities are under no legal obligation to disclose scoring systems to the public generally and to impacted individuals specifically. Some scoring systems are not a secret because their business model is the sale of scores to private and public entities. Data brokers, for instance, rank, categorize, and score consumers on non-credit bases so they can avoid the obligations of FCRA.[160]

To be sure, it is impossible to challenge a scoring system that consumers do not even know exists. Secret scores about people's health, employability, habits, and the like may amount to unfair practices even though they fall outside the requirements of FCRA. In that case, the FTC would have authority to require entities to disclose hidden scoring systems.

[159] Amar Bhide, *The Hidden Costs of Debt Market Liquidity* 17–19 (Ctr. on Capitalism & Soc'y, Columbia Univ., Working Paper No. 79, 2013), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2206996.

[160] Pam Dixon, Exec. Dir., World Privacy Forum, Testimony Before Senate Committee on Commerce Science and Transportation: What Information Do Data Brokers Have On Consumers, and How Do They Use It? 3 (Dec. 18, 2013), *available at* http://www.worldprivacyforum.org/wp-content/uploads/2013/12/WPF_PamDixon_CongressionalTestimony_DataBrokers_2013_fs.pdf. For a discussion of the Fair Credit Reporting Act model, see Frank Pasquale, *Reputation Regulation: Disclosure and the Challenge of Clandestinely Commensurating Computing*, *in* THE OFFENSIVE INTERNET 107, 111–12 (Saul Levmore & Martha C. Nussbaum eds., 2010).



Of course, scoring systems that remain secret would be difficult for the FTC to identify and interrogate. Lawmakers could insist upon the transparency of scoring systems that impact important life opportunities. California, for instance, has been at the forefront of efforts to improve the transparency of businesses' use of consumer information.[161] The FTC has called upon federal lawmakers to pass legislation giving consumers access to the information that data brokers hold about them.[162] In September 2013, Senate Commerce Committee Chairman Jay Rockefeller announced his committee's investigation of the information collection and sharing practices of top data brokers.[163] We are particularly supportive of such efforts—scoring systems can only be meaningfully assessed if they are known and subject to challenge.

CONCLUSION

Imagine a young woman who failed to get a job out of college, and that failure reduced her "employability" score used by potential employers to determine her fitness for work. She found part-time work at a fast food restaurant. Her credit score fell far below 600 without her even knowing it, perhaps because of inferences associated with certain low-paying jobs. Her low credit score *caused* further bad outcomes, cascading into ever more challenging life circumstances. Talent analytics companies categorized her as a "non-innovator" and "waste." With low scores across countless measures, the young woman was unable to get a full-time job.

To quote Wolff and De-Shalit, "without something like the type of action plan set out here, societies are destined to continue to reinforce patterns of entrenched privilege and disadvantage, widening gaps between rich and poor, and perpetuation of disadvantage."[164] Michael Walzer's social theory also provides a compelling argument against the "big data's" promiscuous mashup of various data sources to deny

[161] ACLU OF CAL., LOSING THE SPOTLIGHT: A STUDY OF CALIFORNIA'S SHINE THE LIGHT LAW 13 (2013), *available at* http://www.aclunc.org/R2K.

[162] FED. TRADE COMM'N, PROTECTING CONSUMER PRIVACY IN AN ERA OF RAPID CHANGE 14 (2012), *available at* http://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

[163] Tom Risen, *Rockefeller Expands Investigation of Consumer Data Brokers*, U.S. NEWS & WORLD REP. (Sept. 25, 2013), http://www.usnews.com/news/articles/2013/09/25/rockefeller-expands-investigation-on-consumer-data-brokers.

[164] JONATHAN WOLFF & AVNER DE-SHALIT, DISADVANTAGE 186 (2007).

opportunities.[165] Providing oversight over scoring systems that can cause negative spirals should be a critical aim of our legal system. Scoring systems have a powerful allure—their simplicity gives the illusion of precision and reliability. But predictive algorithms can be anything but accurate and fair. They can narrow people's life opportunities in arbitrary and discriminatory ways.

As a society, we have made commitments to protect consumers from serious harms that they have no means to prevent. We have also aspired to provide individuals with notice about important decisions made about them and a chance to challenge them. These commitments can help us develop a model of due process for scoring systems. Transparency is a crucial first step, first to the FTC who can interrogate scoring systems under their unfairness authority. Opening up the black box scoring systems to individuals or neutral experts representing them is key to permitting them to challenge "arbitrariness by algorithm." Our recommendations are provisional, yet, we hope the FTC and interested lawmakers move forward in bringing procedural regularity and oversight into our scored society.

[165] Mike Konczal, *Demos on Credit Reporting and Employment; Surveillance, Inequalities and the Labor Market*, RORTYBOMB (June 23, 2011), http://rortybomb.wordpress.com/2011/06/23/demos-on-credit-reporting-and-employment-surveillance-inequalities-and-the-labor-market/ ("[Walzer suggested that] nobody should be precluded a social good y because on their lack of possession of an unrelated good x. That the sloppiness of credit scores, the protection of bankruptcy against bad debts, the brute luck of bad health, etc. could all preclude someone from obtaining basic utilities and access to productive labor—that inequality in net worth, health and other spheres preclude access to the sphere of labor regardless of one's abilities—is something to be fought tooth-and-nail.").

Chapter 23

Transparent, Explainable, and Accountable AI for Robotics (Wachter, et al.)

Transparent, Explainable, and Accountable AI for Robotics

Sandra Wachter,[1,2] Brent Mittelstadt,[2,3, 1] Luciano Floridi[1,2]

[1]Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford, OX1 3JS, United Kingdom; [2]The Alan Turing Institute, British Library, 96 Euston Rd, London, NW1 2DB, United Kingdom; [3]Department of Science and Technology Studies, University College London, 22 Gordon Square, London, WC1E 6BT, United Kingdom.

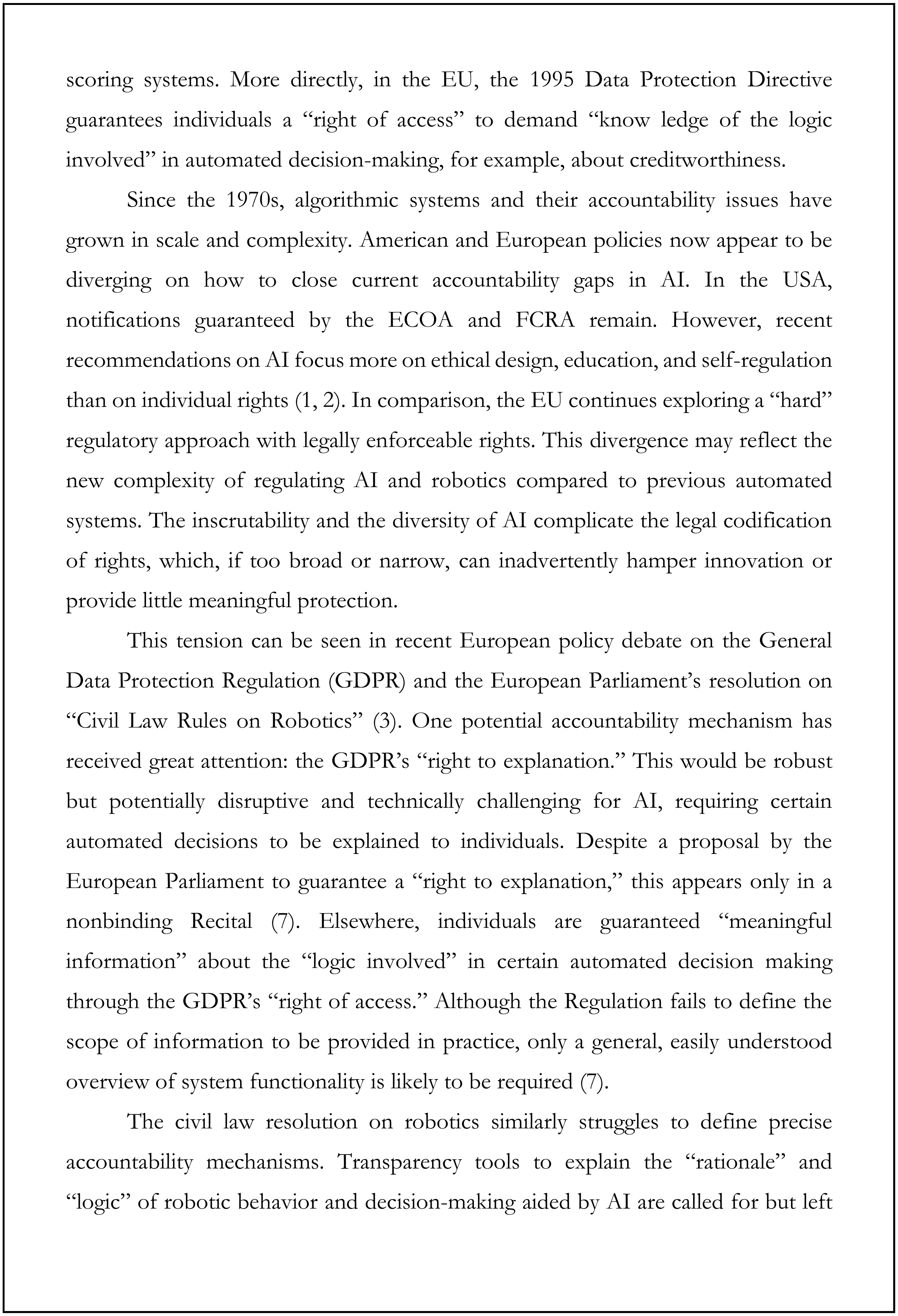
Correspondence: Sandra Wachter, sandra.wachter@oii.ox.ac.uk

Abstract

To create fair and accountable AI and robotics, we need precise regulation and better methods to certify, explain, and audit inscrutable systems.

Recent governmental statements from the United States (USA) (1, 2), the European Union (EU) (3), and China (4) identify artificial intelligence (AI) and robotics as economic and policy priorities. Despite this enthusiasm, challenges remain. Systems can make unfair and discriminatory decisions, replicate or develop biases, and behave in inscrutable and unexpected ways in highly sensitive environments that put human interests and safety at risk (5). For example, Tesla's self-driving cars, policing robot Knightscope, or companion robot Pepper autonomously decides whether something is a pedestrian or another car, whether an individual poses a threat, or which emotion(s) the user is experiencing. In response, pressure is mounting to make algorithms, AI, and robotics fair, trans parent, explainable, and therefore accountable.

These challenges have been reflected in regulation applicable to automated systems since the 1970s. In the USA, the Equal Credit Opportunity Act (ECOA) and the Fair Credit Reporting Act (FCRA) aim to increase transparency in the credit industry (6) and indirectly affect automated systems. Consumers are guaranteed notifications of reasons for adverse actions, including those based on automated



scoring systems. More directly, in the EU, the 1995 Data Protection Directive guarantees individuals a "right of access" to demand "know ledge of the logic involved" in automated decision-making, for example, about creditworthiness.

Since the 1970s, algorithmic systems and their accountability issues have grown in scale and complexity. American and European policies now appear to be diverging on how to close current accountability gaps in AI. In the USA, notifications guaranteed by the ECOA and FCRA remain. However, recent recommendations on AI focus more on ethical design, education, and self-regulation than on individual rights (1, 2). In comparison, the EU continues exploring a "hard" regulatory approach with legally enforceable rights. This divergence may reflect the new complexity of regulating AI and robotics compared to previous automated systems. The inscrutability and the diversity of AI complicate the legal codification of rights, which, if too broad or narrow, can inadvertently hamper innovation or provide little meaningful protection.

This tension can be seen in recent European policy debate on the General Data Protection Regulation (GDPR) and the European Parliament's resolution on "Civil Law Rules on Robotics" (3). One potential accountability mechanism has received great attention: the GDPR's "right to explanation." This would be robust but potentially disruptive and technically challenging for AI, requiring certain automated decisions to be explained to individuals. Despite a proposal by the European Parliament to guarantee a "right to explanation," this appears only in a nonbinding Recital (7). Elsewhere, individuals are guaranteed "meaningful information" about the "logic involved" in certain automated decision making through the GDPR's "right of access." Although the Regulation fails to define the scope of information to be provided in practice, only a general, easily understood overview of system functionality is likely to be required (7).

The civil law resolution on robotics similarly struggles to define precise accountability mechanisms. Transparency tools to explain the "rationale" and "logic" of robotic behavior and decision-making aided by AI are called for but left undefined (3). The Parliament's Committee on Civil Liberties, Justice, and Home Affairs called for compliance with the GDPR in future civil law addressing robotics (8). Several data protection safeguards were explicitly highlighted, including "the right to obtain an explanation" and "information obligations" (e.g., the right of access). Although GDPR compliance is still called for, both safeguards are no longer explicitly mentioned in the final resolution (3). European legislators thus missed a second opportunity to clarify the GDPR's accountability requirements for AI and robotics.

Issues remain, even if future civil law rules for robotics are fully compliant with the GDPR's safeguards against automated decision making. The safeguards only apply to decisions "based solely on automated processing," which may exclude many robotic systems (9). There is reluctance in high-risk areas (e.g., transport) to remove humans entirely from the loop. The outcome may be that robotic decision making would not qualify as "solely" automated. Ironically, this reluctance could make systems less accountable by preventing the GDPR's safeguards from applying. Automated decisions must also have "legal" or "significant" effects for safeguards to apply (Fig. 1), although a definition of such effects is not provided. Only two examples are given: online credit applications and e-recruiting. It remains to be seen whether autonomous robotic behaviors will have "legal" or "significant" effects and how levels of autonomy will influence this definition (9).

Designing imprecise regulation that treats decision-making algorithms, AI, and robotics separately is dangerous. It misinterprets their legal and ethical challenges as unrelated. Concerns about fairness, transparency, interpretability, and accountability are equivalent, have the same genesis, and must be addressed together, regardless of the mix of hardware, software, and data involved. For example, security robots and predictive policing software identify threats with the same method (automated processing) and purpose (public safety). Hence, the desire to understand both systems is the same.

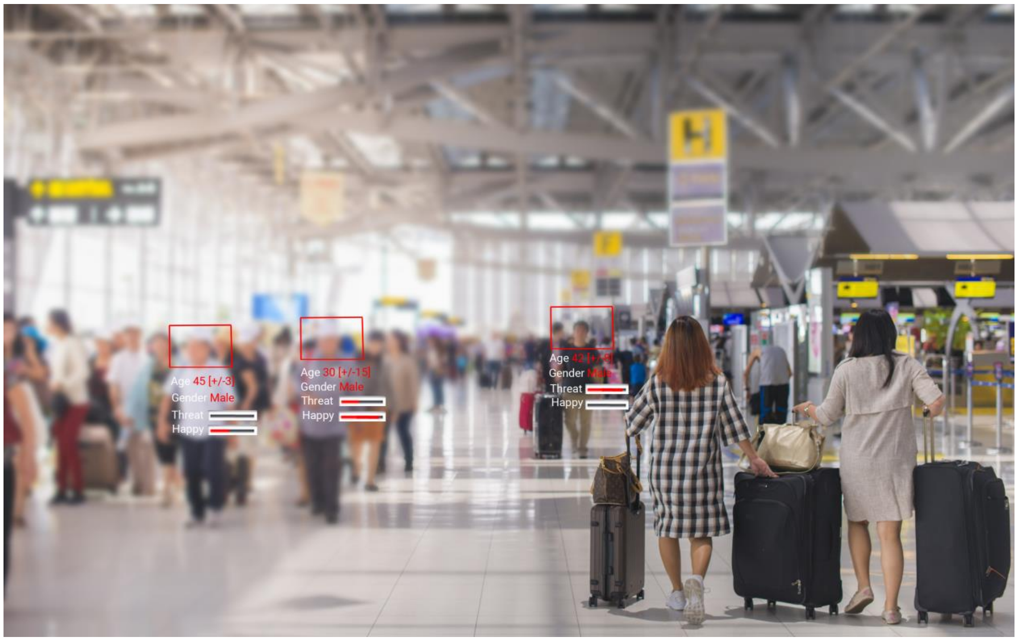


Figure 1 - Security or companion robots detecting threat level or mood solely based on automated processing could produce "significant" effects for an individual, but it remains unclear whether such robotic decisions fall within the scope of the GDPR's safeguards. Photo credit: Shutterstock/Anucha Maneechote. Design: Adham Tamer, Oxford Internet Institute

These issues will only grow in importance. Beijing will soon issue a national development plan for AI (10). It will be interesting to see whether China addresses AI's accountability challenges and, if so, adopts a self-regulatory or "hard law" approach comparable to the USA or EU. Other mechanisms may also be expanded, such as pre-deployment software certification schemes required by China's Cybersecurity Law.

Regulatory and technical accountability mechanisms will be effective only if designed by taking into account the common functionality and diverse complexity of algorithms, AI, and robotics. Several considerations require further research:

- **How can human-interpretable systems be designed without sacrificing performance?** Interpretability is often perceived to be at odds with model accuracy and efficiency in machine learning. In robotics, methods are needed to provide legally required explanations without significantly hampering performance, for example, using proxy or simplified models or rule extraction.

- **How can transparency and accountability be achieved in inscrutable systems?** Inscrutability in AI challenges calls for transparency. Mechanisms not reliant on full interpretability, including pre-deployment certification and algorithmic auditing (5), require further development to ensure transparency and accountability in opaque systems. It remains to be seen whether such "black box" approaches that assess inputs and outputs will comply with legal requirements.
- **How can parallels between emerging systems be identified to set accountability requirements?** Regulatory standards need to be developed to set system- and context-dependent accountability requirements based on potential biased and discriminatory decision-making and risks to safety, fairness, and privacy.

References

1. National Science and Technology Council, "Preparing for the future of artificial intelligence" (Executive Office of the President, 2016), (available at https://www.whitehouse.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf).

2. National Science and Technology Council, "The National Artificial Intelligence Research and Development Strategic Plan" (Executive Office of the President, 2016), (available at https://www.nitrd.gov/PUBS/national_ai_rd_strategic_plan.pdf).

3. European Parliament, "Civil Law Rules on Robotics - European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))" (P8_TA-PROV(2017)00 51, European Parliament, 2017), (available at http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+TA+P8-TA-2017-0051+0+DOC+PDF+V0//EN).

4. X. Bo, China rolls out three-year program for AI growth. *Xinhua News* (2016), (available at http://news.xinhuanet.com/english/2016-05/23/c_135382029.htm).

5. B. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, L. Floridi, The ethics of algorithms: Mapping the debate. *Big Data Soc.* **3** (2016), doi:10.1177/2053951716679679.

6. W. F. Taylor, Meeting the equal credit opportunity act's specificity requirement: Judgmental and statistical scoring systems. *Buffalo Law Rev.* **29**, 73-130 (1980).

7. S. Wachter, B. Mittelstadt, L. Floridi, Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *Int. Data Priv. Law* (2017) (available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903469).

8. European Parliament Committee on Legal Affairs, "Report with recommendations to the Commission on Civil Law Rules on Robotics" (2015/2103(INL), European Parliament, 2017), (available at http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+REPORT+A8-2017-0005+0+DOC+PDF+V0//EN).

9. G.-Z. Yang *et al.*, Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy. *Sci. Robot.* **2**, eaam8638 (2017).

10. M. Jing, Beijing to release national artificial intelligence development plan. *South China Morning Post* (2017), (available at http://www.scmp.com/tech/article/2078209/beijing-release-national-artificial-intelligence-development-plan).

Chapter 24

Rethinking the Mental Steps Doctrine and Other Barriers to Patentability of Artificial Intelligence (Hattenbach and Snyder)

THE COLUMBIA
SCIENCE & TECHNOLOGY
LAW REVIEW

VOL. XIX STLR.ORG SPRING 2018

ARTICLE

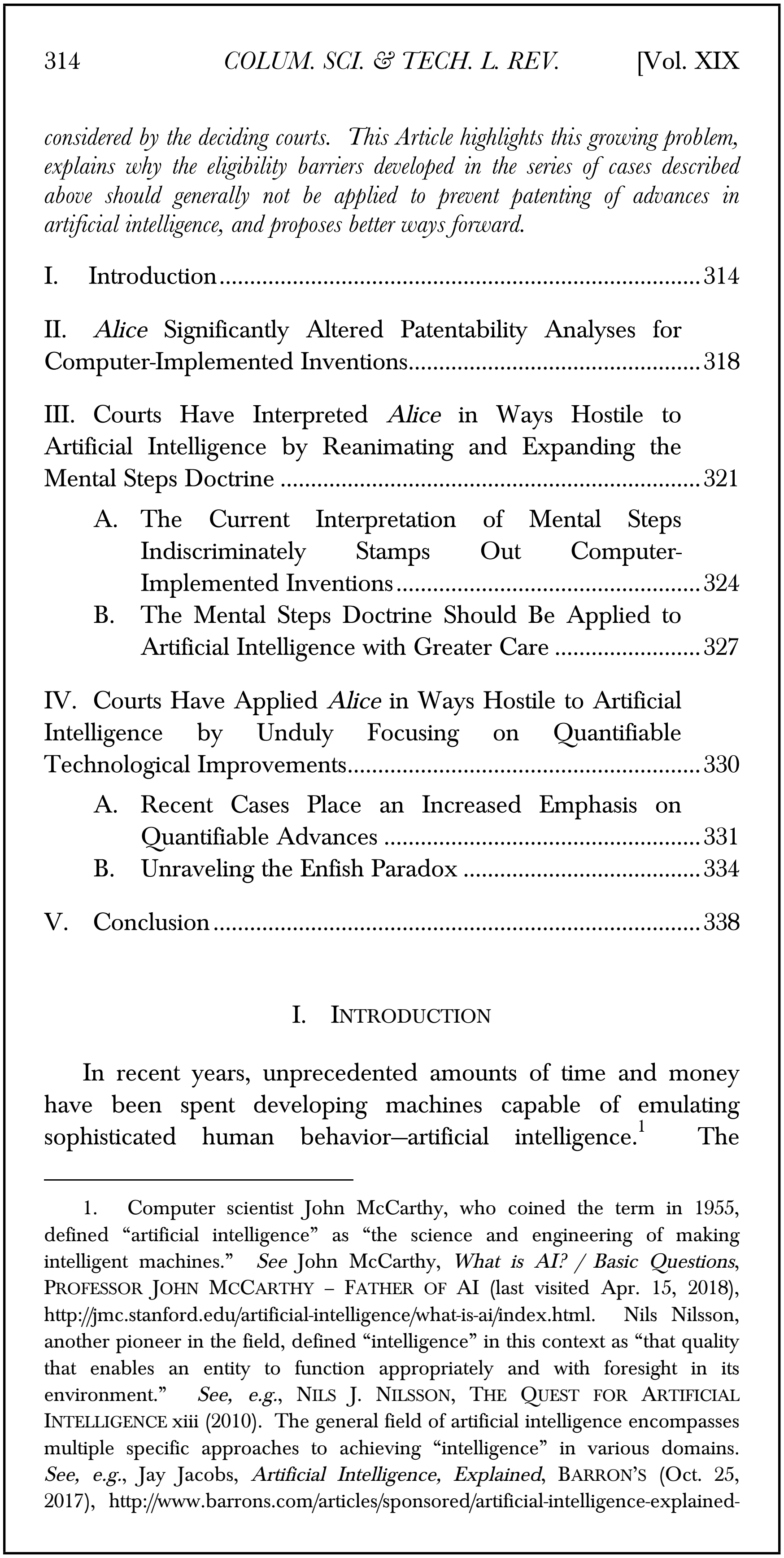
RETHINKING THE MENTAL STEPS DOCTRINE AND OTHER BARRIERS TO PATENTABILITY OF ARTIFICIAL INTELLIGENCE[†]

Ben Hattenbach and Gavin Snyder*

In recent years, our federal courts have given increased attention to the question of what subject matter is eligible for patent protection. The resulting caselaw, developed mostly in the context of business methods or other relatively straight forward technologies, exhibits a number of trends that broadly call into question the patentability of inventions in the field of artificial intelligence. In particular, one series of cases has revitalized the "mental steps doctrine" as a mechanism for invalidating patents. These cases suggest that technology for emulating or replicating activities that could otherwise be accomplished by the human thought process are not patentable. A second series of cases has placed undue emphasis on quantifiable operational improvements as a yardstick for patent-eligibility of computer-related inventions. These cases suggest that even the most ingenious and useful advances in that area may be unpatentable if they do not also provide a readily measurable improvement in performance. Although this precedent was largely crafted outside the artificial intelligence context, it is nonetheless being used by inventors, investors, and courts to gauge the patentability of advances in artificial intelligence. As a result, incentives to innovate in that field are being considerably diminished and, in some instances, altogether eliminated—a consequence that does not appear to have been

[†] This article may be cited as http://www.stlr.org/cite.cgi?volume=19&article=Hattenbach. This work is made available under the Creative Commons Attribution–Non-Commercial–No Derivative Works 3.0 License.

* Mr. Hattenbach is a partner at Irell & Manella LLP and Mr. Snyder is a former associate at Irell & Manella LLP and current associate at Quinn Emanuel Urquhart and Sullivan, LLP. They both specialize in intellectual property litigation. This article does not necessarily represent the views of either firm or their clients.



considered by the deciding courts. This Article highlights this growing problem, explains why the eligibility barriers developed in the series of cases described above should generally not be applied to prevent patenting of advances in artificial intelligence, and proposes better ways forward.

I. INTRODUCTION

In recent years, unprecedented amounts of time and money have been spent developing machines capable of emulating sophisticated human behavior—artificial intelligence.[1] The

[1] Computer scientist John McCarthy, who coined the term in 1955, defined "artificial intelligence" as "the science and engineering of making intelligent machines." *See* John McCarthy, *What is AI? / Basic Questions*, PROFESSOR JOHN MCCARTHY – FATHER OF AI (last visited Apr. 15, 2018), http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html. Nils Nilsson, another pioneer in the field, defined "intelligence" in this context as "that quality that enables an entity to function appropriately and with foresight in its environment." *See, e.g.*, NILS J. NILSSON, THE QUEST FOR ARTIFICIAL INTELLIGENCE xiii (2010). The general field of artificial intelligence encompasses multiple specific approaches to achieving "intelligence" in various domains. *See, e.g.*, Jay Jacobs, *Artificial Intelligence, Explained*, BARRON'S (Oct. 25, 2017), http://www.barrons.com/articles/sponsored/artificial-intelligence-explained-

resulting innovation has already changed life as we know it. Artificial intelligence is being used for pharmaceutical development. Intelligent systems have infiltrated our homes in the form of robotic vacuums and smart thermostats. And they have found their way into our pockets as personal assistants on smartphones. Technologists project that we will have fleets of self-driving cars[2] and affordable domestic robots.[3] We will have devices that flawlessly recognize not only text and speech, but also images. We will have autonomous weapon systems. We will have tools to aid medical determinations in real time, which diagnose illnesses based on an individual patient's genetics and environmental exposure, develop individualized medical treatment plans, and monitor a patient's recovery.[4] Indeed, advances in a variety of fields are already laying the foundations for an "artificial general intelligence" that, like a human, will be able to adapt to many different tasks and environments.[5]

Development of these advances has been and will continue to be quite costly. Companies and governments are investing billions of dollars each year into artificial intelligence research and development.[6] As high as the current costs are, the long-term

1508530169 (techniques to achieve artificial intelligence include machine learning and a specific type of machine learning called deep learning).

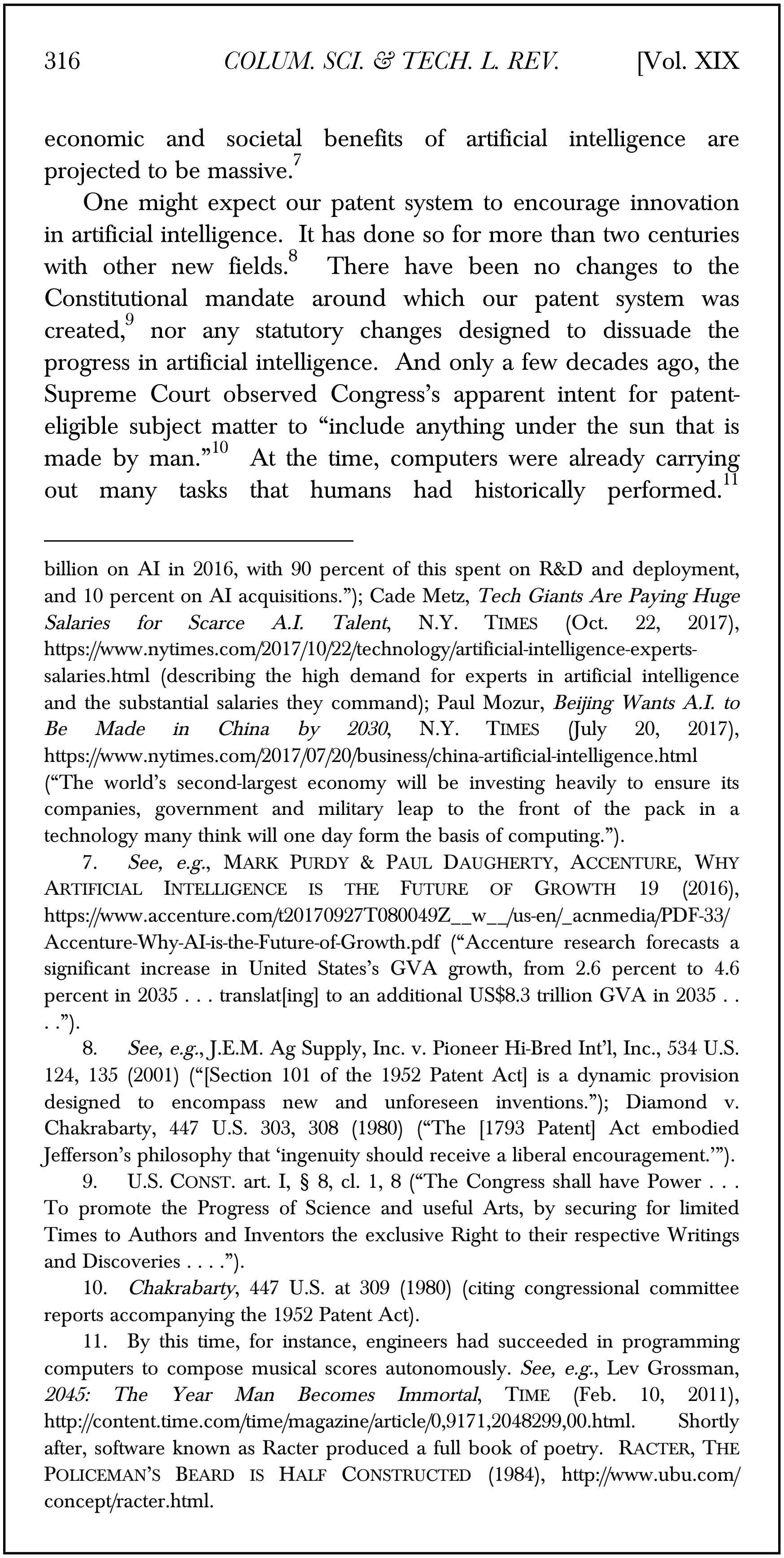
[2] *See, e.g.*, Mike Isaac, *What It Feels Like to Ride in a Self-Driving Uber*, N.Y. TIMES (Sept. 14, 2016), http://www.nytimes.com/2016/09/15/technology/our-reporter-goes-for-a-spin-in-a-self-driving-uber-car.html (describing Uber's September 2016 pilot test of a small number of driverless cars in Pittsburgh).

[3] *See, e.g.*, Sharon Gaudin, *Elon Musk Wants to Build You a Robotic Housekeeper*, COMPUTERWORLD (June 21, 2016, 12:55 PM), https://www.computerworld.com/article/3086931/artificial-intelligence/elon-musk-wants-to-build-you-a-robotic-housekeeper.html.

[4] *See, e.g.,* Yuichi Mori et al., *Computer-Aided Diagnosis for Colonoscopy*, 49 ENDOSCOPY 813 (2017), https://www.researchgate.net/publication/317147787_Computer-aided_diagnosis_for_colonoscopy (using machine learning to automatically detect and classify potentially cancerous polyps during colonoscopy).

[5] Gideon Lewis-Kraus, *The Great A.I. Awakening*, N.Y. TIMES MAG. (Dec. 14, 2016), https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html ("Artificial general intelligence will not involve dutiful adherence to explicit instructions, but instead will demonstrate a facility with the implicit, the interpretive. It will be a general tool, designed for general purposes in a general context.").

[6] *See, e.g.*, JACQUES BUGHIN ET AL., MCKINSEY & CO., ARTIFICIAL INTELLIGENCE: THE NEXT DIGITAL FRONTIER? 4 (June 2017), https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/how-artificial-intelligence-can-deliver-real-value-to-companies (follow "Discussion Paper" hyperlink) ("Globally, we estimate tech giants spent $20 billion to $30



economic and societal benefits of artificial intelligence are projected to be massive.[7]

One might expect our patent system to encourage innovation in artificial intelligence. It has done so for more than two centuries with other new fields.[8] There have been no changes to the Constitutional mandate around which our patent system was created,[9] nor any statutory changes designed to dissuade the progress in artificial intelligence. And only a few decades ago, the Supreme Court observed Congress's apparent intent for patent-eligible subject matter to "include anything under the sun that is made by man."[10] At the time, computers were already carrying out many tasks that humans had historically performed.[11]

billion on AI in 2016, with 90 percent of this spent on R&D and deployment, and 10 percent on AI acquisitions."); Cade Metz, *Tech Giants Are Paying Huge Salaries for Scarce A.I. Talent*, N.Y. TIMES (Oct. 22, 2017), https://www.nytimes.com/2017/10/22/technology/artificial-intelligence-experts-salaries.html (describing the high demand for experts in artificial intelligence and the substantial salaries they command); Paul Mozur, *Beijing Wants A.I. to Be Made in China by 2030*, N.Y. TIMES (July 20, 2017), https://www.nytimes.com/2017/07/20/business/china-artificial-intelligence.html ("The world's second-largest economy will be investing heavily to ensure its companies, government and military leap to the front of the pack in a technology many think will one day form the basis of computing.").

[7] *See, e.g.*, MARK PURDY & PAUL DAUGHERTY, ACCENTURE, WHY ARTIFICIAL INTELLIGENCE IS THE FUTURE OF GROWTH 19 (2016), https://www.accenture.com/t20170927T080049Z__w__/us-en/_acnmedia/PDF-33/Accenture-Why-AI-is-the-Future-of-Growth.pdf ("Accenture research forecasts a significant increase in United States's GVA growth, from 2.6 percent to 4.6 percent in 2035 . . . translat[ing] to an additional US$8.3 trillion GVA in 2035 . . . .").

[8] *See, e.g.*, J.E.M. Ag Supply, Inc. v. Pioneer Hi-Bred Int'l, Inc., 534 U.S. 124, 135 (2001) ("[Section 101 of the 1952 Patent Act] is a dynamic provision designed to encompass new and unforeseen inventions."); Diamond v. Chakrabarty, 447 U.S. 303, 308 (1980) ("The [1793 Patent] Act embodied Jefferson's philosophy that 'ingenuity should receive a liberal encouragement.'").

[9] U.S. CONST. art. I, § 8, cl. 1, 8 ("The Congress shall have Power . . . To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries . . . .").

[10] *Chakrabarty*, 447 U.S. at 309 (1980) (citing congressional committee reports accompanying the 1952 Patent Act).

[11] By this time, for instance, engineers had succeeded in programming computers to compose musical scores autonomously. *See, e.g.*, Lev Grossman, *2045: The Year Man Becomes Immortal*, TIME (Feb. 10, 2011), http://content.time.com/time/magazine/article/0,9171,2048299,00.html. Shortly after, software known as Racter produced a full book of poetry. RACTER, THE POLICEMAN'S BEARD IS HALF CONSTRUCTED (1984), http://www.ubu.com/concept/racter.html.

Everyone knew that computers would continue to become increasingly sophisticated and, along the way, take over an expanding array of functions from humans. People welcomed these developments, recognizing that they would lead to vast improvements in quality of life. Indeed, artificial intelligence advances were, and are, widely celebrated in both the marketplace[12] and in popular culture.[13]

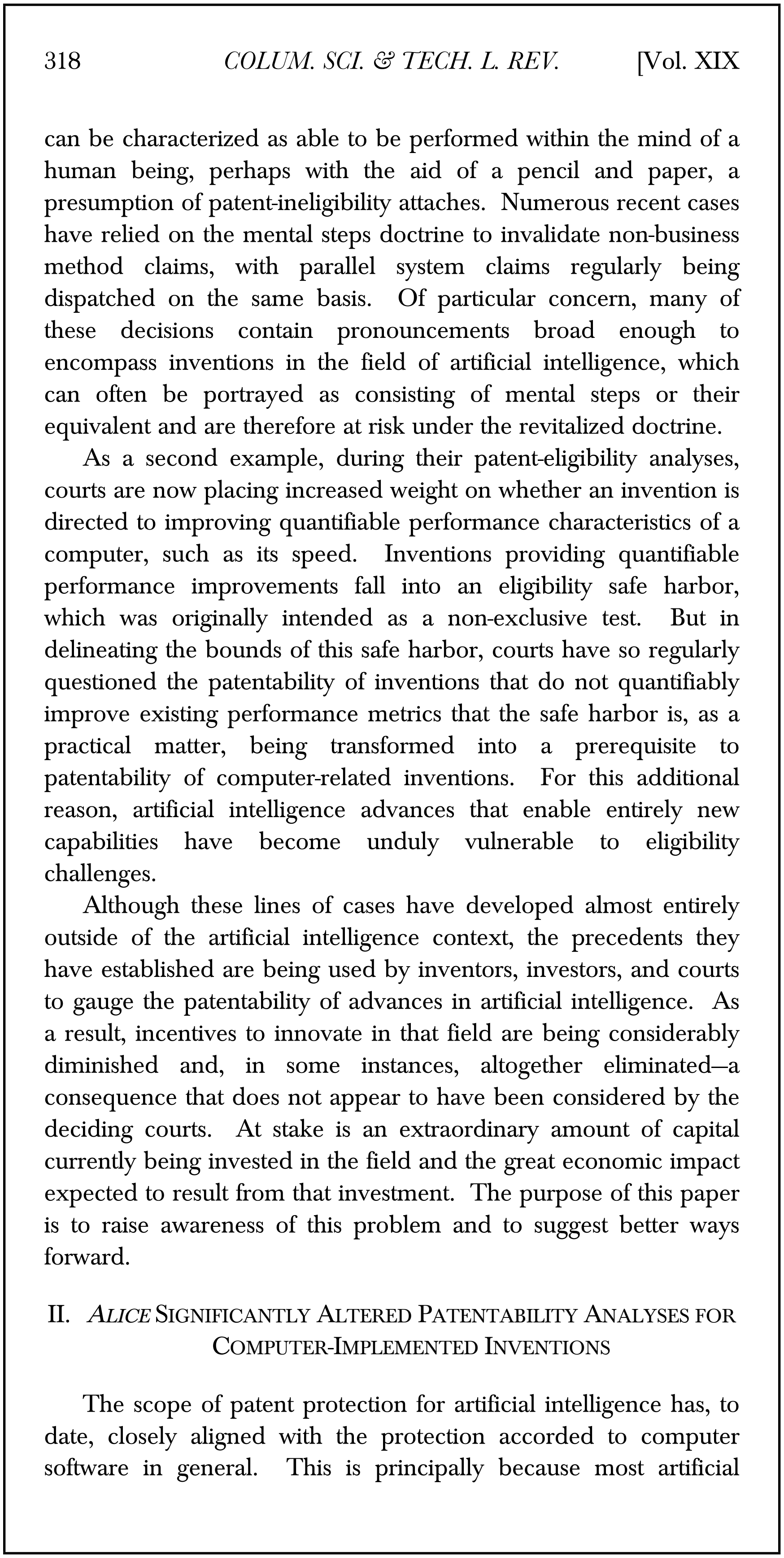
There has nonetheless been a recent sea change in the scope of patent protection for artificial intelligence. Following *Alice Corp. v. CLS Bank International*,[14] many courts have delivered broad pronouncements on the scope of patentable subject matter. In many instances, these edicts have been rendered in the context of pure business methods or quintessentially abstract ideas undeserving of patent protection. But unfortunately, many of the opinions are filled with expansive language that, if taken at face value, extend well beyond the circumstances of the cases being decided and into vastly dissimilar fields. If these opinions are removed from their original contexts and applied indiscriminately to the field of artificial intelligence, they would severely curtail or even eliminate patent protection for legitimate inventions.

As one example of this, courts in the aftermath of *Alice* have revived the "mental steps" doctrine as a primary yardstick for assessing patent-eligibility. Under this doctrine, if method claims

[12] The services of many highly valued companies, such as Google and Facebook, depend heavily on automated recognition of what humans find relevant. *See, e.g.*, Vindu Goel, *When Yahoo Ruled the Valley: Stories of the Original 'Surfers'*, N.Y. TIMES (July 16, 2016), http://www.nytimes.com/2016/07/17/technology/when-yahoo-ruled-the-valley-stories-of-the-original-surfers.html (discussing the early use of humans to catalog information on the web as having been "long since eclipsed by Google and Facebook"); Gil Press, *Why Yahoo Lost and Google Won*, FORBES (July 26, 2016), http://www.forbes.com/sites/gilpress/2016/07/26/why-yahoo-lost-and-google-won/ (citing automation as "the heart of Google's success").

[13] For example, artificially intelligent computer systems were featured in multiple episodes of the original *Star Trek* series (1966–69), including one titled "The Ultimate Computer," in which a computer capable of learning and adapting was given command of the Enterprise. The television show *Knight Rider* (1982–86) focused on the exploits of a futuristic, computer-driven sports car run using artificial intelligence. *Buck Rogers in the 25th Century* (1979–1981) involved an array of artificially intelligent robots that served as assistants to humans. 2001: A SPACE ODYSSEY (1968) and its sequel 2010: THE YEAR WE MAKE CONTACT (1984) both centered on artificial intelligence-based computer systems. THE TERMINATOR (1984) featured a human-looking cyborg sent from the future. Perhaps most famously, the original STAR WARS trilogy (1977–1983) included robots that had human-like capabilities and foibles.

[14] 134 S. Ct. 2347 (2014).



can be characterized as able to be performed within the mind of a human being, perhaps with the aid of a pencil and paper, a presumption of patent-ineligibility attaches. Numerous recent cases have relied on the mental steps doctrine to invalidate non-business method claims, with parallel system claims regularly being dispatched on the same basis. Of particular concern, many of these decisions contain pronouncements broad enough to encompass inventions in the field of artificial intelligence, which can often be portrayed as consisting of mental steps or their equivalent and are therefore at risk under the revitalized doctrine.

As a second example, during their patent-eligibility analyses, courts are now placing increased weight on whether an invention is directed to improving quantifiable performance characteristics of a computer, such as its speed. Inventions providing quantifiable performance improvements fall into an eligibility safe harbor, which was originally intended as a non-exclusive test. But in delineating the bounds of this safe harbor, courts have so regularly questioned the patentability of inventions that do not quantifiably improve existing performance metrics that the safe harbor is, as a practical matter, being transformed into a prerequisite to patentability of computer-related inventions. For this additional reason, artificial intelligence advances that enable entirely new capabilities have become unduly vulnerable to eligibility challenges.

Although these lines of cases have developed almost entirely outside of the artificial intelligence context, the precedents they have established are being used by inventors, investors, and courts to gauge the patentability of advances in artificial intelligence. As a result, incentives to innovate in that field are being considerably diminished and, in some instances, altogether eliminated—a consequence that does not appear to have been considered by the deciding courts. At stake is an extraordinary amount of capital currently being invested in the field and the great economic impact expected to result from that investment. The purpose of this paper is to raise awareness of this problem and to suggest better ways forward.

II. *Alice* Significantly Altered Patentability Analyses for Computer-Implemented Inventions

The scope of patent protection for artificial intelligence has, to date, closely aligned with the protection accorded to computer software in general. This is principally because most artificial intelligence innovations, at least historically, *are* software.[15] Many potential advances in artificial intelligence will include algorithms capable of executing on a standard computer or smartphone. The scope has also aligned because no judicially-recognized distinctions between software in general, and artificial intelligence software in particular, have yet arisen. Such distinctions were previously unnecessary because for many years nearly all varieties of software, except for the most abstractly mathematical, were considered patent-eligible by default. A famous line of Supreme Court cases, ending with *Diamond v. Diehr*, had generally supported the patentability of computer software throughout the 1980s, 1990s, and 2000s.[16] Software that had some practical application in the real world was protected.[17]

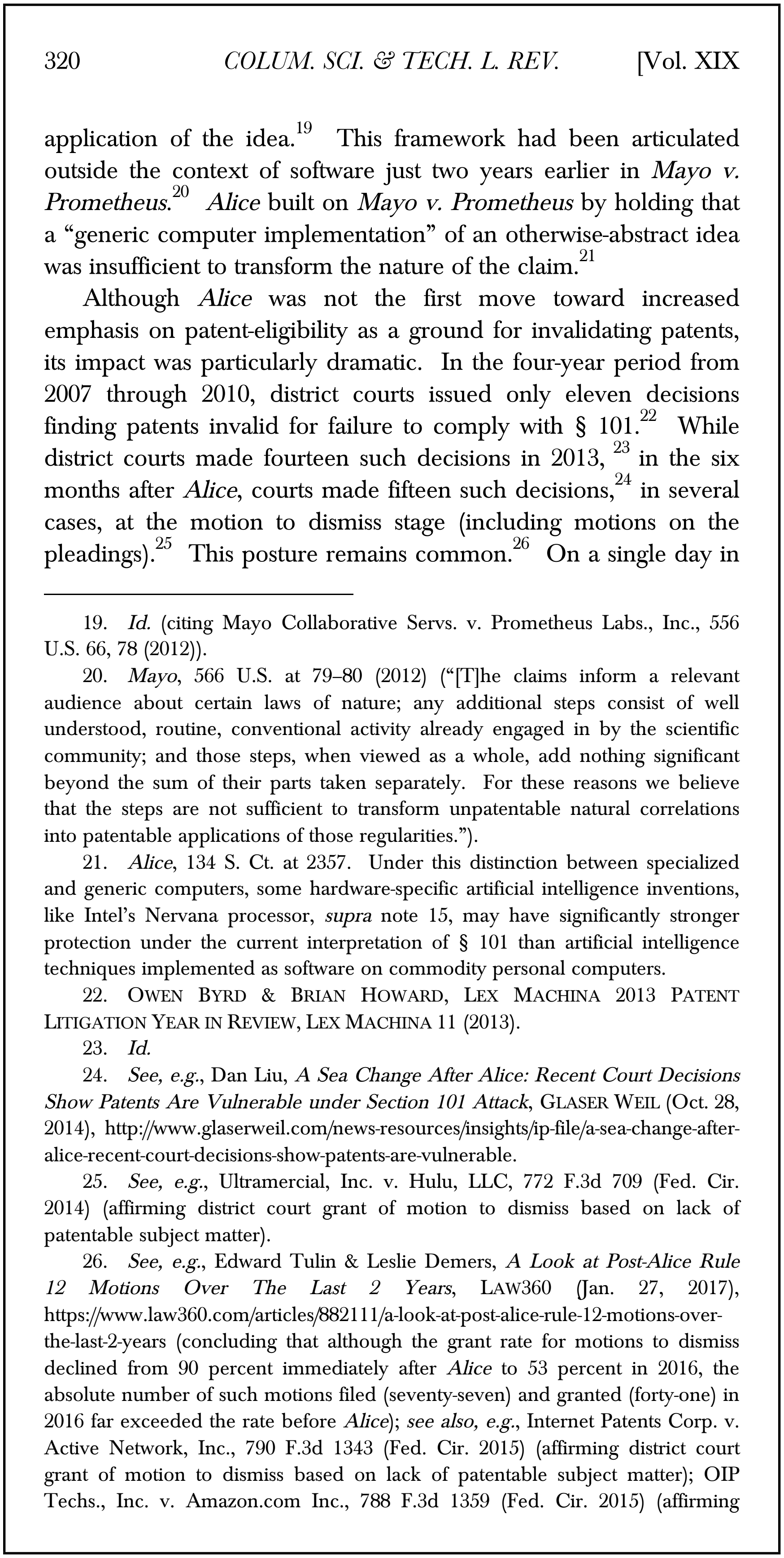
The Supreme Court's most recent general guidance on patent-eligibility in *Alice* applied a two-step framework in which a court first assesses whether a claim is directed to a "patent-ineligible concept."[18] If so, the court next asks whether any of the claim's elements "transform the nature of the claim" into a patent-eligible

[15] There are, of course, artificial intelligence inventions that use specialized, new hardware. For example, Intel recently debuted "the world's first family of processors designed from the ground up for artificial intelligence (AI)." Naveen Rao, *Intel Nervana Neural Network Processors (NNP) Redefine AI Silicon*, INTEL AI (Oct. 17, 2017), https://ai.intel.com/intel-nervana-neural-network-processors-nnp-redefine-ai-silicon/.

[16] *See, e.g.,* Gottschalk v. Benson, 409 U.S. 63, 71–72 (1972) (holding patent-ineligible an algorithm for performing certain numerical conversions on a general-purpose digital computer) ("It is said that the decision precludes a patent for any program servicing a computer. We do not so hold."); Parker v. Flook, 437 U.S. 584, 584, 590 (1978) (holding that a computer-implemented algorithm is not made patent-eligible by "identification of a limited category of useful, though conventional, post-solution applications," but noting that "a process is not unpatentable simply because it contains . . . a mathematical algorithm."); Diamond v. Diehr, 450 U.S. 175, 187–88 (1981) (holding that using a well-known mathematical equation in a real-world rubber curing process was patent-eligible and noting that "a claim drawn to subject matter otherwise statutory does not become nonstatutory simply because it uses a mathematical formula, computer program, or digital computer").

[17] *See, e.g.*, Arrhythmia Research Tech., Inc. v. Corazonix Corp., 958 F.2d 1053, 1056–57 (Fed. Cir. 1992) (finding algorithm for computer analysis of electrocardiographic signals to determine heart function patentable subject matter) ("As the jurisprudence developed, inventions that were implemented by the mathematically-directed performance of computers were viewed in the context of the practical application to which the computer-generated data were put.").

[18] Alice Corp. v. CLS Bank Int'l, 134 S. Ct. 2347, 2355 (2014) (referring to the three historical categories of patent-ineligible subject matter under §101: laws of nature, natural phenomena, and abstract ideas).



application of the idea.[19] This framework had been articulated outside the context of software just two years earlier in *Mayo v. Prometheus*.[20] *Alice* built on *Mayo v. Prometheus* by holding that a "generic computer implementation" of an otherwise-abstract idea was insufficient to transform the nature of the claim.[21]

Although *Alice* was not the first move toward increased emphasis on patent-eligibility as a ground for invalidating patents, its impact was particularly dramatic. In the four-year period from 2007 through 2010, district courts issued only eleven decisions finding patents invalid for failure to comply with § 101.[22] While district courts made fourteen such decisions in 2013, [23] in the six months after *Alice*, courts made fifteen such decisions,[24] in several cases, at the motion to dismiss stage (including motions on the pleadings).[25] This posture remains common.[26] On a single day in

[19] *Id.* (citing Mayo Collaborative Servs. v. Prometheus Labs., Inc., 556 U.S. 66, 78 (2012)).

[20] *Mayo*, 566 U.S. at 79–80 (2012) ("[T]he claims inform a relevant audience about certain laws of nature; any additional steps consist of well understood, routine, conventional activity already engaged in by the scientific community; and those steps, when viewed as a whole, add nothing significant beyond the sum of their parts taken separately. For these reasons we believe that the steps are not sufficient to transform unpatentable natural correlations into patentable applications of those regularities.").

[21] *Alice*, 134 S. Ct. at 2357. Under this distinction between specialized and generic computers, some hardware-specific artificial intelligence inventions, like Intel's Nervana processor, *supra* note 15, may have significantly stronger protection under the current interpretation of § 101 than artificial intelligence techniques implemented as software on commodity personal computers.

[22] OWEN BYRD & BRIAN HOWARD, LEX MACHINA 2013 PATENT LITIGATION YEAR IN REVIEW, LEX MACHINA 11 (2013).

[23] *Id.*

[24] *See, e.g.*, Dan Liu, *A Sea Change After Alice: Recent Court Decisions Show Patents Are Vulnerable under Section 101 Attack*, GLASER WEIL (Oct. 28, 2014), http://www.glaserweil.com/news-resources/insights/ip-file/a-sea-change-after-alice-recent-court-decisions-show-patents-are-vulnerable.

[25] *See, e.g.*, Ultramercial, Inc. v. Hulu, LLC, 772 F.3d 709 (Fed. Cir. 2014) (affirming district court grant of motion to dismiss based on lack of patentable subject matter).

[26] *See, e.g.*, Edward Tulin & Leslie Demers, *A Look at Post-Alice Rule 12 Motions Over The Last 2 Years*, LAW360 (Jan. 27, 2017), https://www.law360.com/articles/882111/a-look-at-post-alice-rule-12-motions-over-the-last-2-years (concluding that although the grant rate for motions to dismiss declined from 90 percent immediately after *Alice* to 53 percent in 2016, the absolute number of such motions filed (seventy-seven) and granted (forty-one) in 2016 far exceeded the rate before *Alice*); *see also, e.g.*, Internet Patents Corp. v. Active Network, Inc., 790 F.3d 1343 (Fed. Cir. 2015) (affirming district court grant of motion to dismiss based on lack of patentable subject matter); OIP Techs., Inc. v. Amazon.com Inc., 788 F.3d 1359 (Fed. Cir. 2015) (affirming

September 2014, five different decisions invalidated software patents under *Alice*.[27] There were as many or more invalidations under § 101 on that one day than in any single *year* between 2007 and 2011.[28] Since *Alice*, claims of more than 500 separate patents have been found invalid under § 101.[29]

The lack of an explicit definition of an "abstract idea" in *Alice* itself has led the lower courts to rule primarily by analogy to the facts of previous cases.[30] While each decision applying *Alice* is the response of a court to particular circumstances, general themes have emerged from lower courts' attempts to distill and apply the Supreme Court's guidance. Some of these themes are particularly troublesome for artificial intelligence.

III. Courts Have Interpreted *Alice* in Ways Hostile to Artificial Intelligence by Reanimating and Expanding the Mental Steps Doctrine

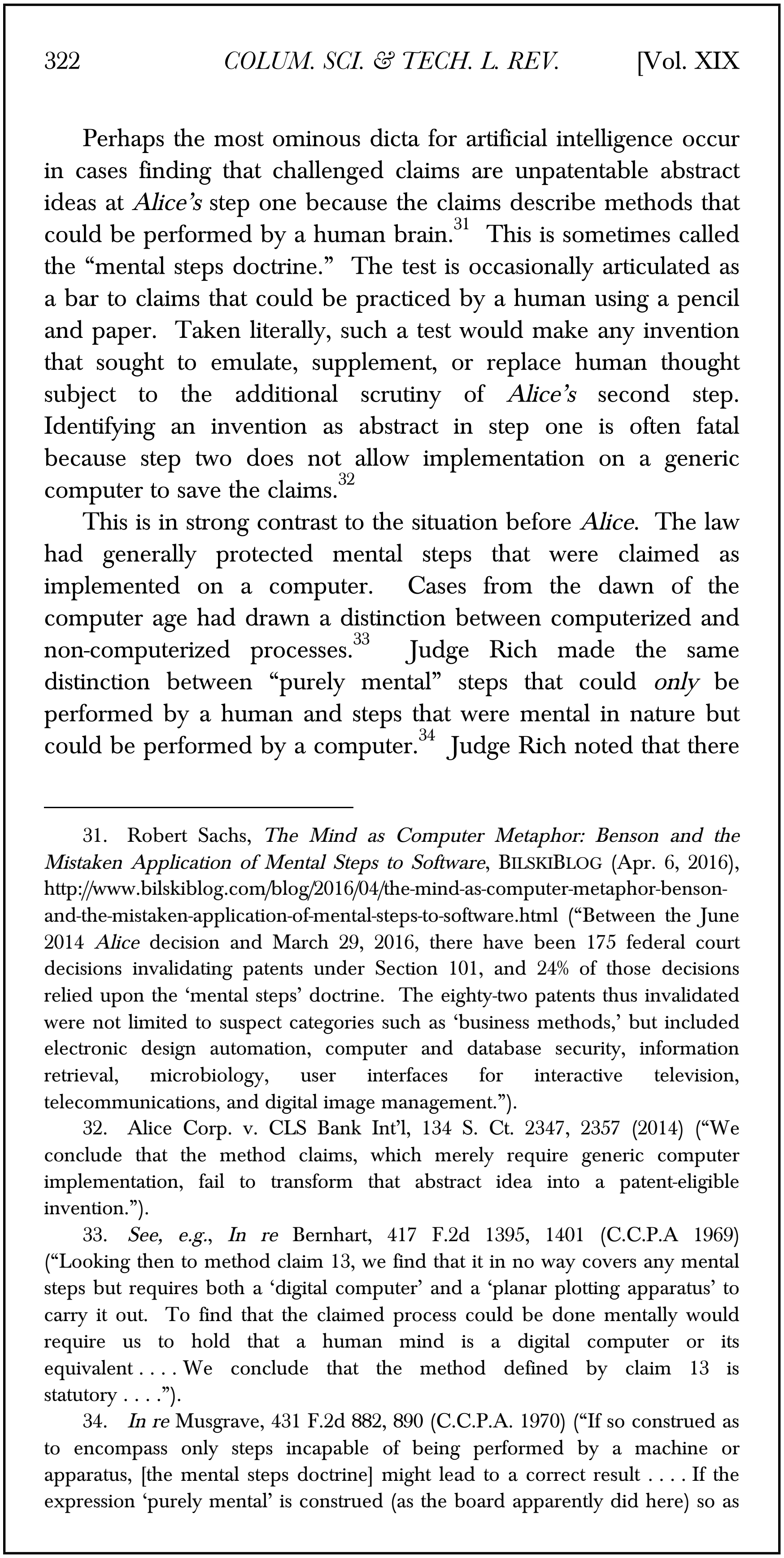
district court grant of motion for judgment on the pleadings based on lack of patentable subject matter).

[27] *See, e.g.*, Gregory Garre, et al., *Early Lessons on Alice Corp. v. CLS Bank International and Section 101 From Recent Court Decisions*, LATHAM & WATKINS (Sept. 19, 2014), *available at* http://www.lw.com/thoughtLeadership/lw-alice-corp-cls-bank-section-101.

[28] BYRD & HOWARD, *supra* note 22, at 11 (discovering that no more than five patents were invalidated for lack of patentable subject matter between the years of 2007–2011).

[29] Robert Sachs, *#Alicestorm: April Update and the Impact of TC Heartland on Patent Eligibility*, BILSKIBLOG (June 1, 2017), http://www.bilskiblog.com/blog/2017/06/alicestorm-april-update-and-the-impact-of-tc-heartland.html (stating that, as of June 1, 2017, district courts had invalidated claims of 515 patents under § 101 in 242 decisions since *Alice* and the Federal Circuit had affirmed 91.7% of the appeals from those decisions).

[30] *See, e.g.*, Amdocs (Israel) Ltd. v. Openet Telecom, Inc., 841 F.3d 1288, 1294 (Fed. Cir. 2016) ("Whether the more detailed analysis is undertaken at step one or at step two, the analysis presumably would be based on a generally-accepted and understood *definition* of, or test for, what an 'abstract idea' encompasses. However, a search for a single test or definition in the decided cases concerning § 101 from this court, and indeed from the Supreme Court, reveals that at present there is no such single, succinct, usable definition or test. The problem with articulating a single, universal definition of 'abstract idea' is that it is difficult to fashion a workable definition to be applied to as-yet-unknown cases with as-yet-unknown inventions. That is not for want of trying; to the extent the efforts so far have been unsuccessful it is because they often end up using alternative but equally abstract terms or are overly narrow. Instead of a definition, then, the decisional mechanism courts now apply is to examine earlier cases in which a similar or parallel descriptive nature can be seen—what prior cases were about, and which way they were decided.").



Perhaps the most ominous dicta for artificial intelligence occur in cases finding that challenged claims are unpatentable abstract ideas at *Alice's* step one because the claims describe methods that could be performed by a human brain.[31] This is sometimes called the "mental steps doctrine." The test is occasionally articulated as a bar to claims that could be practiced by a human using a pencil and paper. Taken literally, such a test would make any invention that sought to emulate, supplement, or replace human thought subject to the additional scrutiny of *Alice's* second step. Identifying an invention as abstract in step one is often fatal because step two does not allow implementation on a generic computer to save the claims.[32]

This is in strong contrast to the situation before *Alice*. The law had generally protected mental steps that were claimed as implemented on a computer. Cases from the dawn of the computer age had drawn a distinction between computerized and non-computerized processes.[33] Judge Rich made the same distinction between "purely mental" steps that could *only* be performed by a human and steps that were mental in nature but could be performed by a computer.[34] Judge Rich noted that there

[31] Robert Sachs, *The Mind as Computer Metaphor: Benson and the Mistaken Application of Mental Steps to Software*, BilskiBlog (Apr. 6, 2016), http://www.bilskiblog.com/blog/2016/04/the-mind-as-computer-metaphor-benson-and-the-mistaken-application-of-mental-steps-to-software.html ("Between the June 2014 *Alice* decision and March 29, 2016, there have been 175 federal court decisions invalidating patents under Section 101, and 24% of those decisions relied upon the 'mental steps' doctrine. The eighty-two patents thus invalidated were not limited to suspect categories such as 'business methods,' but included electronic design automation, computer and database security, information retrieval, microbiology, user interfaces for interactive television, telecommunications, and digital image management.").

[32] Alice Corp. v. CLS Bank Int'l, 134 S. Ct. 2347, 2357 (2014) ("We conclude that the method claims, which merely require generic computer implementation, fail to transform that abstract idea into a patent-eligible invention.").

[33] *See, e.g.*, *In re* Bernhart, 417 F.2d 1395, 1401 (C.C.P.A 1969) ("Looking then to method claim 13, we find that it in no way covers any mental steps but requires both a 'digital computer' and a 'planar plotting apparatus' to carry it out. To find that the claimed process could be done mentally would require us to hold that a human mind is a digital computer or its equivalent . . . . We conclude that the method defined by claim 13 is statutory . . . .").

[34] *In re* Musgrave, 431 F.2d 882, 890 (C.C.P.A. 1970) ("If so construed as to encompass only steps incapable of being performed by a machine or apparatus, [the mental steps doctrine] might lead to a correct result . . . . If the expression 'purely mental' is construed (as the board apparently did here) so as

was no support in the statute for barring mental steps performed by a computer.[35] Indeed, there were many such claims that covered new technological improvements that deserved patent protection.[36] Although ultimately invalidating the challenged patent, dicta in *Parker v. Flook* similarly appeared to reject the "pencil and paper" test if the claimed process was primarily intended to be computerized.[37]

Although the Federal Circuit in *In re Comiskey* criticized "mental processes standing alone" as unpatentable abstract ideas,[38] that case was limited to claims that covered mental steps that were *not* computerized.[39] The mental steps doctrine was also used before *Alice* to invalidate claims for which the goal was not to replace or augment human activity, but simply to cover a medical diagnostic test that involves a human comparison between two or

to encompass steps performable by apparatus, as well as mentally, then the [doctrine] is unsound . . . .").

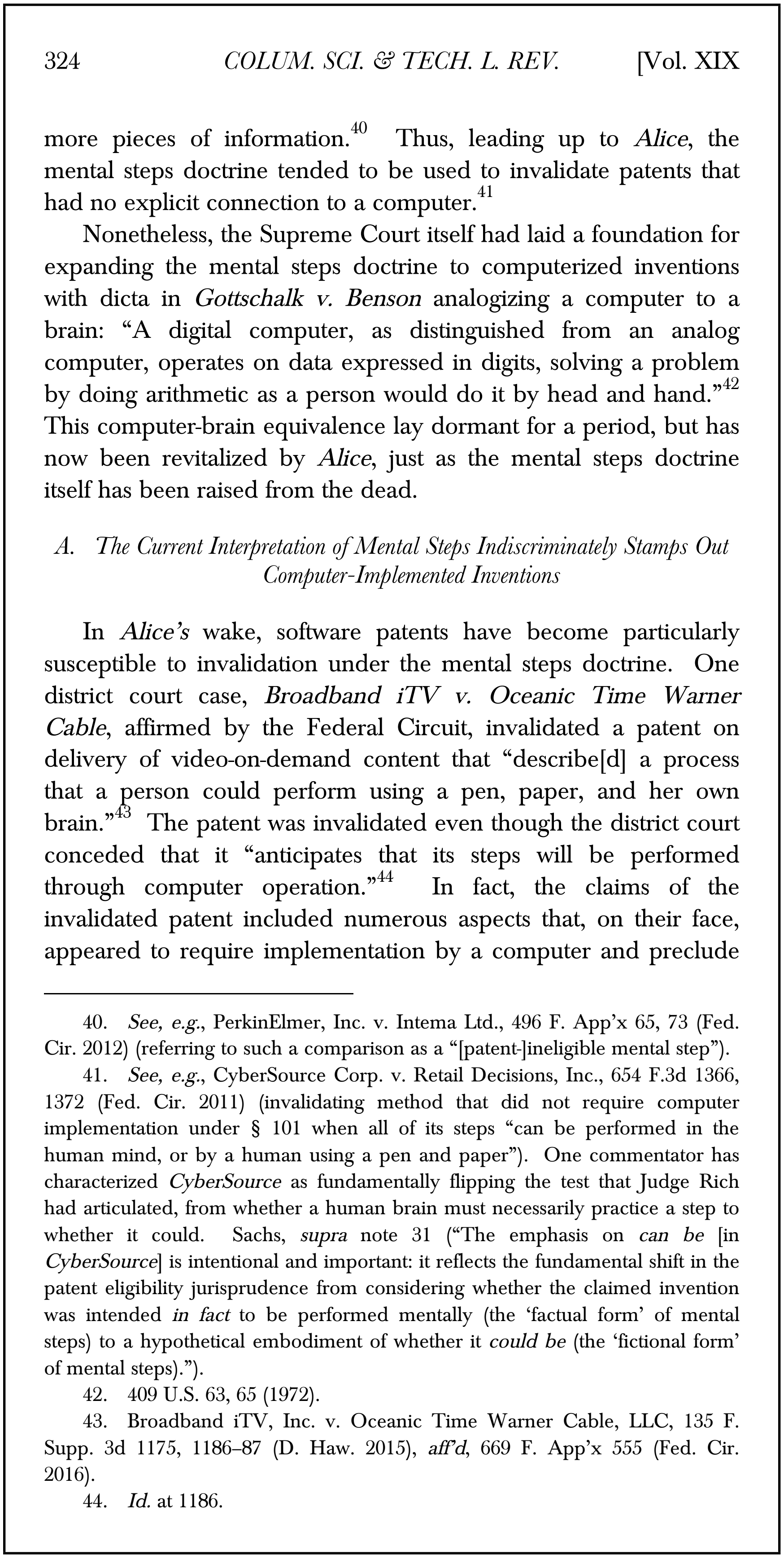
[35] *Id.* ("As may be seen from the statutory language, it contains nothing whatever which would either include or exclude claims containing 'mental steps' and whatever law there may be on the subject cannot be attributed to Congress.").

[36] *Id.* at 893 ("We cannot agree with the board that these claims (all the steps of which can be carried out by the disclosed apparatus) are directed to non-statutory processes merely because some or all of the steps therein can also be carried out in or with the aid of the human mind or because it may be necessary for one performing the processes to think. All that is necessary, in our view, to make a sequence of operational steps a statutory 'process' within 35 U.S.C. § 101 is that it be in the technological arts so as to be in consonance with the Constitutional purpose to promote the progress of 'useful arts.'" (quoting U.S. CONST. art. I, § 8, cl. 8)).

[37] 437 U.S. 584, 586 (1978) ("Although the computations can be made by pencil and paper calculations, the abstract of disclosure makes it clear that the formula is primarily useful for computerized calculations producing automatic adjustments in alarm settings.").

[38] *In re* Comiskey, 554 F.3d 967, 979 (Fed Cir. 2009) ("[M]ental processes—or processes of human thinking—standing alone are not patentable even if they have practical application.").

[39] *Id.* at 980 ("It is thus clear that the present statute does not allow patents to be issued on particular business systems—such as a particular type of arbitration—that depend *entirely* on the use of mental processes. In other words, the patent statute does not allow patents on particular systems that depend for their operation *on human intelligence alone*, a field of endeavor that both the framers and Congress intended to be beyond the reach of patentable subject matter. Thus, it is established that the application of human intelligence to the solution of practical problems is not *in and of itself* patentable.") (emphasis added); *id.* at 970 ("[T]he parties agree that these claims do not require . . . the use of a mechanical device such as a computer.").



more pieces of information.[40] Thus, leading up to *Alice*, the mental steps doctrine tended to be used to invalidate patents that had no explicit connection to a computer.[41]

Nonetheless, the Supreme Court itself had laid a foundation for expanding the mental steps doctrine to computerized inventions with dicta in *Gottschalk v. Benson* analogizing a computer to a brain: "A digital computer, as distinguished from an analog computer, operates on data expressed in digits, solving a problem by doing arithmetic as a person would do it by head and hand."[42] This computer-brain equivalence lay dormant for a period, but has now been revitalized by *Alice*, just as the mental steps doctrine itself has been raised from the dead.

A. *The Current Interpretation of Mental Steps Indiscriminately Stamps Out Computer-Implemented Inventions*

In *Alice's* wake, software patents have become particularly susceptible to invalidation under the mental steps doctrine. One district court case, *Broadband iTV v. Oceanic Time Warner Cable*, affirmed by the Federal Circuit, invalidated a patent on delivery of video-on-demand content that "describe[d] a process that a person could perform using a pen, paper, and her own brain."[43] The patent was invalidated even though the district court conceded that it "anticipates that its steps will be performed through computer operation."[44] In fact, the claims of the invalidated patent included numerous aspects that, on their face, appeared to require implementation by a computer and preclude

[40] *See, e.g.*, PerkinElmer, Inc. v. Intema Ltd., 496 F. App'x 65, 73 (Fed. Cir. 2012) (referring to such a comparison as a "[patent-]ineligible mental step").

[41] *See, e.g.*, CyberSource Corp. v. Retail Decisions, Inc., 654 F.3d 1366, 1372 (Fed. Cir. 2011) (invalidating method that did not require computer implementation under § 101 when all of its steps "can be performed in the human mind, or by a human using a pen and paper"). One commentator has characterized *CyberSource* as fundamentally flipping the test that Judge Rich had articulated, from whether a human brain must necessarily practice a step to whether it could. Sachs, *supra* note 31 ("The emphasis on *can be* [in *CyberSource*] is intentional and important: it reflects the fundamental shift in the patent eligibility jurisprudence from considering whether the claimed invention was intended *in fact* to be performed mentally (the 'factual form' of mental steps) to a hypothetical embodiment of whether it *could be* (the 'fictional form' of mental steps).").

[42] 409 U.S. 63, 65 (1972).

[43] Broadband iTV, Inc. v. Oceanic Time Warner Cable, LLC, 135 F. Supp. 3d 1175, 1186–87 (D. Haw. 2015), *aff'd*, 669 F. App'x 555 (Fed. Cir. 2016).

[44] *Id.* at 1186.

implementation by a human alone, including enabling the "uploading [of] video content in a digital video format via an online network" and "converting the content uploaded to the Web-based content management server into a standard TV digital format."[45] The district court discounted these additional details in its analysis of *Alice* step two, finding that no meaningful additional ingredients were added to the abstract mental steps.[46]

Another recent Federal Circuit case, *Coffelt v. NVIDIA*, invalidated claims directed to "deriving a pixel color in a graphic image."[47] Even though the claims explicitly required a computer to perform various algorithm steps, the court in effect rewrote the claims to eliminate the computer requirement and substituted a more abstract version of the actual claim language that was fitted to mental performance by a human.[48] This rewrite of the claims was justified on the basis that the claimed computer was general-purpose and thus squarely addressed by *Alice*'s instructions on how to apply step two.[49]

Yet another contemporary Federal Circuit case, *FairWarning v. Iatric*, upheld the district court's invalidation of claims directed to computer-implemented fraud detection techniques.[50] The

[45] U.S. Patent No. 7,631,336 at col. 2 l. 60–66 (filed Mar. 12, 2007).

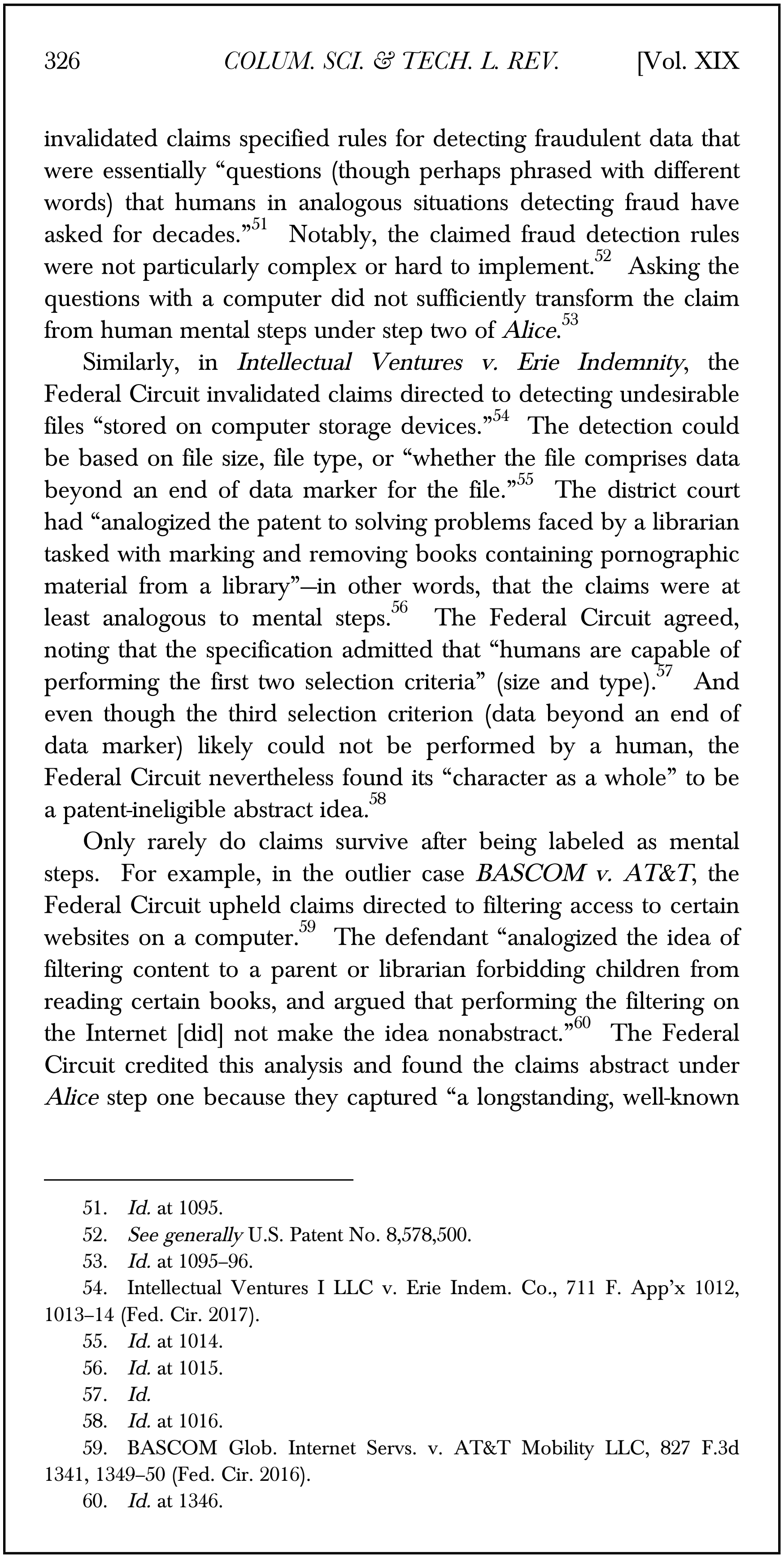
[46] *Broadband iTV*, 135 F. Supp. 3d at 1190 ("Moreover, the fact that a patent provides specific details of implementation is not enough to secure patent eligibility if those 'details' continue to encompass merely 'generic computer implementation' and 'routine activities.'"). The district court also repeated troubling dicta about how it would not matter for patent purposes if the inventor were the first person to implement the process on a computer. *Id.* at 1187 ("[T]he fact that a company may be the first to successfully apply an abstract idea within a new technological context does not transform the abstract idea into something tangible and patentable.") (quoting OpenTV, Inc. v. Apple, Inc., Case No. 14-cv-01622-HSG, 2015 WL 1535328, at *6 (N.D. Cal. Apr. 6, 2015)).

[47] Coffelt v. NVIDIA Corp., 680 F. App'x 1010 (Fed. Cir. 2017).

[48] U.S. Patent No. 8,614,710 at col. 14 l. 1–3 ("for a first pixel, a computer deriving a pixel color for said first position vector from a result of said length comparison"); *Coffelt*, 680 F. App'x at 1011 ("[T]he claims at issue here are directed to the abstract idea of calculating and comparing regions in space . . . . The claims thus recite nothing more than a mathematical algorithm that could be implemented using a pen and paper.").

[49] *Coffelt*, 680 F. App'x at 1011 ("The parties do not dispute that the claims can be implemented on a generic computer . . . . [T]he inventive concept must 'transform' the patent-ineligible algorithm into a 'patent-eligible application' of the algorithm, and do so by more than merely implementing the algorithm on a generic computer.") (citing Alice Corp. v. CLS Bank Int'l, 134 S. Ct. 2347, 2355 (2014)).

[50] FairWarning IP, LLC v. Iatric Sys., Inc., 839 F.3d 1089, 1094–95 (Fed. Cir. 2016).



invalidated claims specified rules for detecting fraudulent data that were essentially "questions (though perhaps phrased with different words) that humans in analogous situations detecting fraud have asked for decades."[51] Notably, the claimed fraud detection rules were not particularly complex or hard to implement.[52] Asking the questions with a computer did not sufficiently transform the claim from human mental steps under step two of *Alice*.[53]

Similarly, in *Intellectual Ventures v. Erie Indemnity*, the Federal Circuit invalidated claims directed to detecting undesirable files "stored on computer storage devices."[54] The detection could be based on file size, file type, or "whether the file comprises data beyond an end of data marker for the file."[55] The district court had "analogized the patent to solving problems faced by a librarian tasked with marking and removing books containing pornographic material from a library"—in other words, that the claims were at least analogous to mental steps.[56] The Federal Circuit agreed, noting that the specification admitted that "humans are capable of performing the first two selection criteria" (size and type).[57] And even though the third selection criterion (data beyond an end of data marker) likely could not be performed by a human, the Federal Circuit nevertheless found its "character as a whole" to be a patent-ineligible abstract idea.[58]

Only rarely do claims survive after being labeled as mental steps. For example, in the outlier case *BASCOM v. AT&T*, the Federal Circuit upheld claims directed to filtering access to certain websites on a computer.[59] The defendant "analogized the idea of filtering content to a parent or librarian forbidding children from reading certain books, and argued that performing the filtering on the Internet [did] not make the idea nonabstract."[60] The Federal Circuit credited this analysis and found the claims abstract under *Alice* step one because they captured "a longstanding, well-known

[51] *Id.* at 1095.

[52] *See generally* U.S. Patent No. 8,578,500.

[53] *Id.* at 1095–96.

[54] Intellectual Ventures I LLC v. Erie Indem. Co., 711 F. App'x 1012, 1013–14 (Fed. Cir. 2017).

[55] *Id.* at 1014.

[56] *Id.* at 1015.

[57] *Id.*

[58] *Id.* at 1016.

[59] BASCOM Glob. Internet Servs. v. AT&T Mobility LLC, 827 F.3d 1341, 1349–50 (Fed. Cir. 2016).

[60] *Id.* at 1346.

method of organizing human behavior."[61] The claims were nevertheless saved under *Alice* step two because they recited "a specific, discrete implementation of the abstract idea of filtering content" and because the patent described "how its particular arrangement of elements [was] a technical improvement over prior art ways of filtering such content."[62]

The cases exhibit some common themes, each potentially hazardous when applied to artificial intelligence. First, a majority of computer-implemented processes, including those underlying much of artificial intelligence, could probably be characterized as mental steps by a judge interpreting that doctrine expansively. There are many judicial descriptions of what does or does not qualify as mental steps that, if applied broadly in the artificial intelligence context, would make patenting in the area quite difficult. For example, the Federal Circuit has suggested that method claims that are merely "the equivalent of human mental work . . . are unpatentable abstract ideas."[63] Second, little to no weight is being given to claim elements that explicitly require computerization. For example, the Federal Circuit has made general pronouncements that "abstract ideas are essentially mental steps; they are not tangible even if they are written down or programmed into a physical machine."[64] Third, even computer-implemented *system* claims, which are manifestly not directed to mental steps, have nonetheless been treated as though they recited mental steps and invalidated where they were perceived as similar to other claims that did qualify as mental steps.[65]

B. The Mental Steps Doctrine Should Be Applied to Artificial Intelligence with Greater Care

We submit that particular caution should be taken with the mental steps doctrine in the context of artificial intelligence

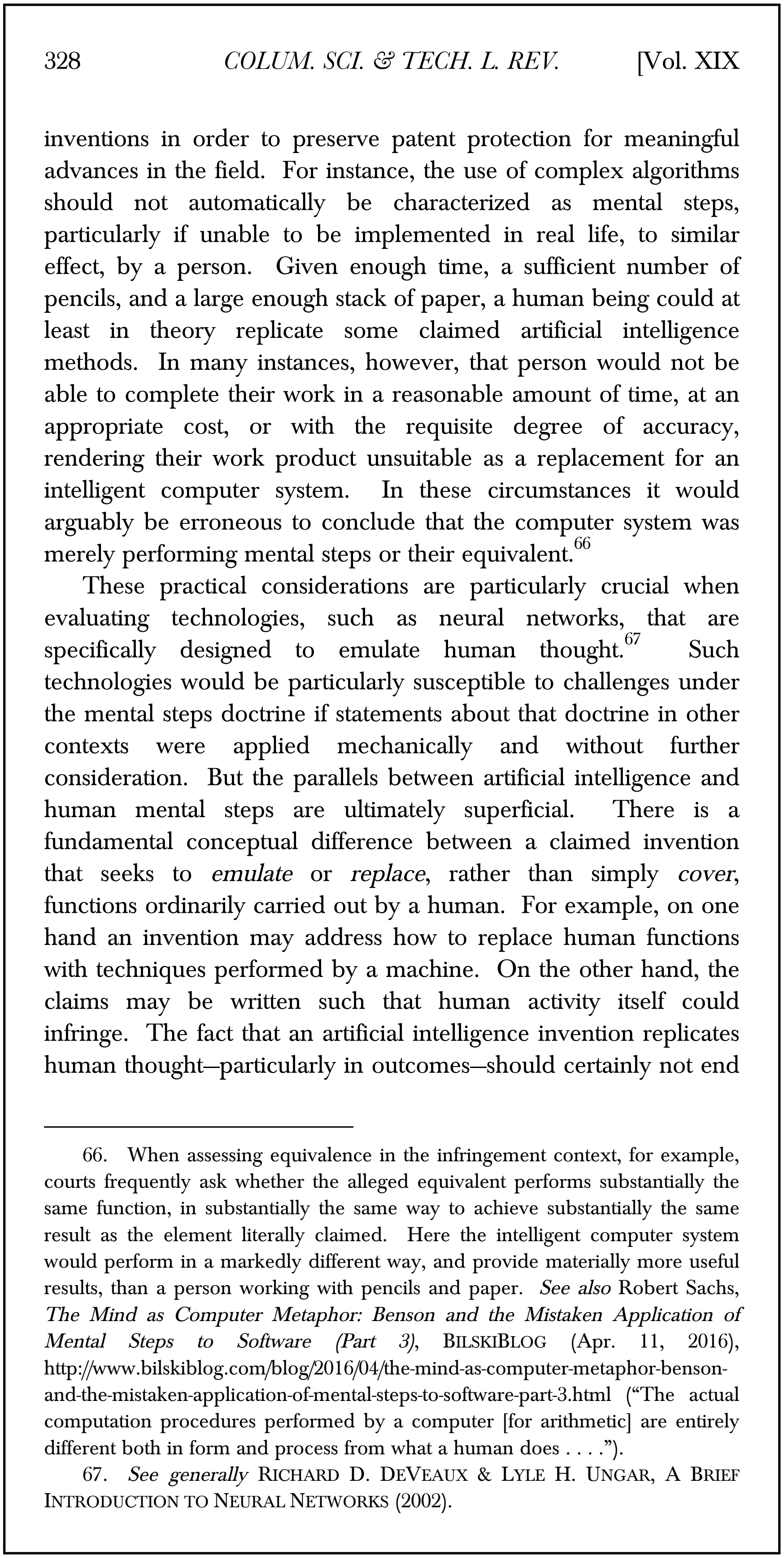
[61] *Id.* at 1348.

[62] *Id.* at 1350.

[63] CyberSource Corp. v. Retail Decisions, Inc., 654 F.3d 1366, 1371 (Fed. Cir. 2011).

[64] Ariosa Diagnostics, Inc. v. Sequenom, Inc., 809 F.3d 1282, 1285 (Fed. Cir. 2015); *see also* Elec. Power Grp., LLC v. Alstom S.A., 830 F.3d 1350, 1355 (Fed. Cir. 2016) ("In a similar vein, we have treated analyzing information by steps people go through in their minds, *or by mathematical algorithms*, without more, as essentially mental processes within the abstract-idea category.") (emphasis added).

[65] *See, e.g.*, Intellectual Ventures I LLC v. Erie Indem. Co., 711 F. App'x 1012 (Fed. Cir. 2017) (invalidating all claims of U.S. Patent No. 7,757,298, including system claim 10, but only providing analysis of method claim 1).



inventions in order to preserve patent protection for meaningful advances in the field. For instance, the use of complex algorithms should not automatically be characterized as mental steps, particularly if unable to be implemented in real life, to similar effect, by a person. Given enough time, a sufficient number of pencils, and a large enough stack of paper, a human being could at least in theory replicate some claimed artificial intelligence methods. In many instances, however, that person would not be able to complete their work in a reasonable amount of time, at an appropriate cost, or with the requisite degree of accuracy, rendering their work product unsuitable as a replacement for an intelligent computer system. In these circumstances it would arguably be erroneous to conclude that the computer system was merely performing mental steps or their equivalent.[66]

These practical considerations are particularly crucial when evaluating technologies, such as neural networks, that are specifically designed to emulate human thought.[67] Such technologies would be particularly susceptible to challenges under the mental steps doctrine if statements about that doctrine in other contexts were applied mechanically and without further consideration. But the parallels between artificial intelligence and human mental steps are ultimately superficial. There is a fundamental conceptual difference between a claimed invention that seeks to *emulate* or *replace*, rather than simply *cover*, functions ordinarily carried out by a human. For example, on one hand an invention may address how to replace human functions with techniques performed by a machine. On the other hand, the claims may be written such that human activity itself could infringe. The fact that an artificial intelligence invention replicates human thought—particularly in outcomes—should certainly not end

[66] When assessing equivalence in the infringement context, for example, courts frequently ask whether the alleged equivalent performs substantially the same function, in substantially the same way to achieve substantially the same result as the element literally claimed. Here the intelligent computer system would perform in a markedly different way, and provide materially more useful results, than a person working with pencils and paper. *See also* Robert Sachs, *The Mind as Computer Metaphor: Benson and the Mistaken Application of Mental Steps to Software (Part 3)*, BILSKIBLOG (Apr. 11, 2016), http://www.bilskiblog.com/blog/2016/04/the-mind-as-computer-metaphor-benson-and-the-mistaken-application-of-mental-steps-to-software-part-3.html ("The actual computation procedures performed by a computer [for arithmetic] are entirely different both in form and process from what a human does . . . .").

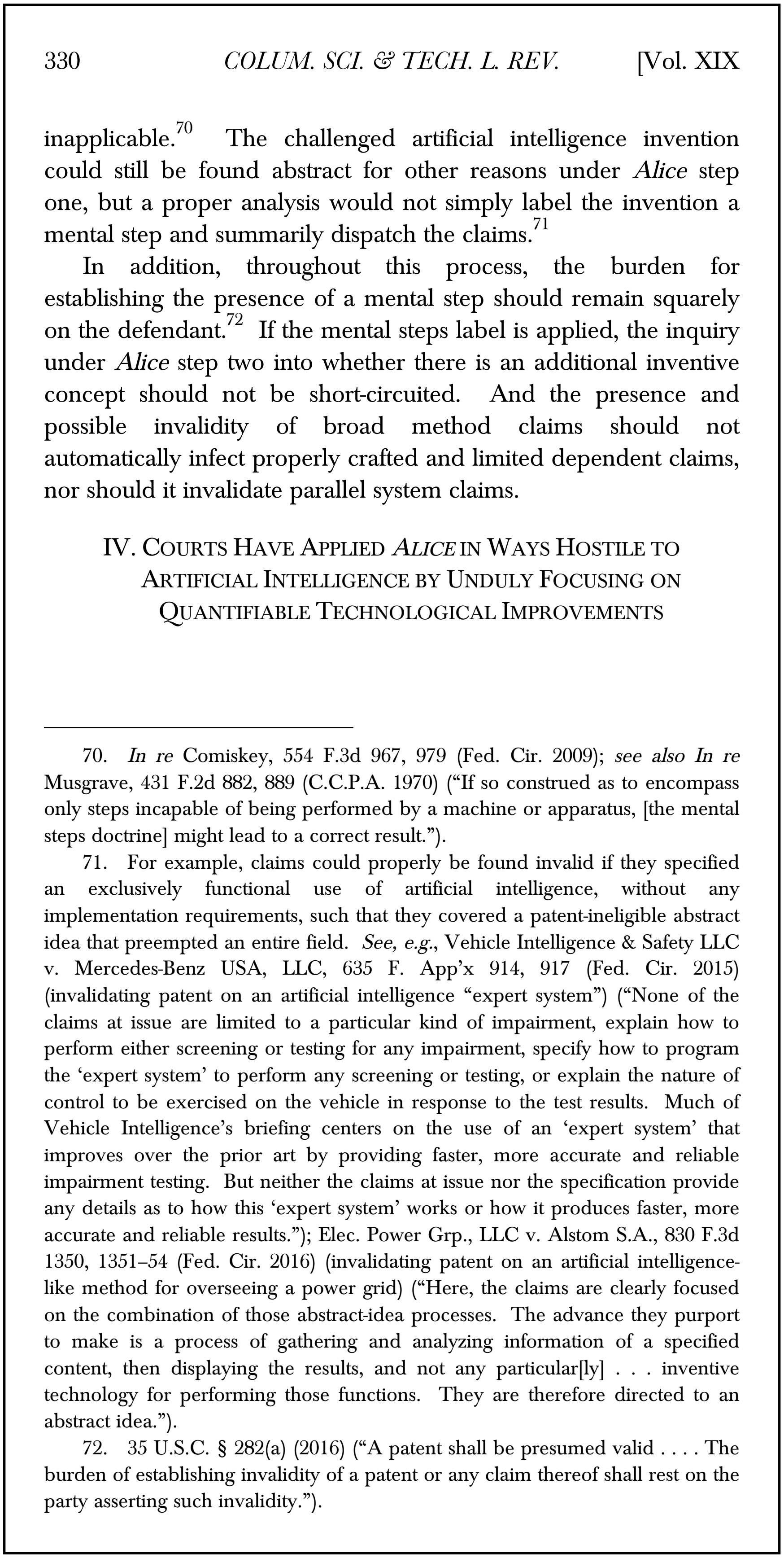
[67] *See generally* RICHARD D. DEVEAUX & LYLE H. UNGAR, A BRIEF INTRODUCTION TO NEURAL NETWORKS (2002).

the patentability analysis; indeed, it arguably should not even be a factor weighing against patent-eligibility. The inquiry should instead focus on the extent to which the challenged claims improperly extend beyond computation or mechanization to cover exclusively human activity.[68]

Furthermore, an analysis of these issues should be conducted from start to finish with attention to the specific requirements of the challenged claims. Meaningful limitations requiring computation or mechanization should not be read out of the claims when the presence of a mental step is being evaluated. For instance, courts should respect claim limitations that explicitly require computation and should avoid sweeping pronouncements that implementation on a "general-purpose computer" is not entitled to any weight in the patentability analysis. *Alice* did not go that far. Arguably, the most that *Alice* holds is that implementation of an otherwise abstract idea on a generic computer cannot save a claim at *Alice* step two.[69] As applied to artificial intelligence claims, nothing in *Alice* is inconsistent with, at *Alice* step one, applying the approach from *In re Comiskey* where an explicit claim limitation requiring computation by a machine makes the mental steps doctrine

[68] Nonetheless, cases that apply the mental steps doctrine frequently seem to object to the appearance of a claim that could make ordinary human activity infringing. *See, e.g.*, *generally*, PerkinElmer, Inc. v. Intema Ltd., 496 F. App'x 65 (Fed. Cir. 2012); Synopsys Inc. v. Mentor Graphics Corp., 839 F.3d 1138 (Fed. Cir. 2017) ("A review of the actual claims at issue shows that they are directed to the abstract idea of translating a functional description of a logic circuit into a hardware component description of the logic circuit. This idea of reviewing a description of certain functions and turning it into a representation of the logic component that performs those functions can be—and, indeed, was—performed mentally or by pencil and paper by one of ordinary skill in the art. Moreover, the claims do not call for the involvement of a computer."); *id.* at 1149 ("On their face, the claims do not call for any form of computer implementation of the claimed methods . . . . Because the Asserted Claims make no mention of employing a computer or any other physical device, they are so broad as to read on an individual performing the claimed steps mentally or with pencil and paper."). *But see, e.g.*, FairWarning IP, LLC v. Iatric Sys. Inc., 839 F.3d 1089, 1094–95 (Fed. Cir. 2016) (invalidating claims that required a computer but were analogous to human mental steps).

[69] Alice Corp. v. CLS Bank Int'l, 134 S. Ct. 2347, 2351 (2014) ("Here, the representative method claim does no more than simply instruct the practitioner to implement the abstract idea of intermediated settlement on a generic computer."); *see* also Gottschalk v. Benson, 409 U.S. 63, 64–65 (1972) (finding that a mathematical algorithm run on a "general purpose digital computer" is not patentable).



inapplicable.[70] The challenged artificial intelligence invention could still be found abstract for other reasons under *Alice* step one, but a proper analysis would not simply label the invention a mental step and summarily dispatch the claims.[71]

In addition, throughout this process, the burden for establishing the presence of a mental step should remain squarely on the defendant.[72] If the mental steps label is applied, the inquiry under *Alice* step two into whether there is an additional inventive concept should not be short-circuited. And the presence and possible invalidity of broad method claims should not automatically infect properly crafted and limited dependent claims, nor should it invalidate parallel system claims.

IV. Courts Have Applied *Alice* in Ways Hostile to Artificial Intelligence by Unduly Focusing on Quantifiable Technological Improvements

[70] *In re* Comiskey, 554 F.3d 967, 979 (Fed. Cir. 2009); *see also In re* Musgrave, 431 F.2d 882, 889 (C.C.P.A. 1970) ("If so construed as to encompass only steps incapable of being performed by a machine or apparatus, [the mental steps doctrine] might lead to a correct result.").

[71] For example, claims could properly be found invalid if they specified an exclusively functional use of artificial intelligence, without any implementation requirements, such that they covered a patent-ineligible abstract idea that preempted an entire field. *See, e.g.*, Vehicle Intelligence & Safety LLC v. Mercedes-Benz USA, LLC, 635 F. App'x 914, 917 (Fed. Cir. 2015) (invalidating patent on an artificial intelligence "expert system") ("None of the claims at issue are limited to a particular kind of impairment, explain how to perform either screening or testing for any impairment, specify how to program the 'expert system' to perform any screening or testing, or explain the nature of control to be exercised on the vehicle in response to the test results. Much of Vehicle Intelligence's briefing centers on the use of an 'expert system' that improves over the prior art by providing faster, more accurate and reliable impairment testing. But neither the claims at issue nor the specification provide any details as to how this 'expert system' works or how it produces faster, more accurate and reliable results."); Elec. Power Grp., LLC v. Alstom S.A., 830 F.3d 1350, 1351–54 (Fed. Cir. 2016) (invalidating patent on an artificial intelligence-like method for overseeing a power grid) ("Here, the claims are clearly focused on the combination of those abstract-idea processes. The advance they purport to make is a process of gathering and analyzing information of a specified content, then displaying the results, and not any particular[ly] . . . inventive technology for performing those functions. They are therefore directed to an abstract idea.").

[72] 35 U.S.C. § 282(a) (2016) ("A patent shall be presumed valid . . . . The burden of establishing invalidity of a patent or any claim thereof shall rest on the party asserting such invalidity.").

There is a second major trend in the post-*Alice* caselaw that should concern those seeking to protect artificial intelligence inventions: courts are now placing excessive emphasis during the patent-eligibility analysis on whether an invention improves traditional computer performance metrics, such as speed or memory capacity. Although such considerations were originally introduced to define a safe harbor protecting certain types of inventions, that safe harbor has in recent years become so central to analysis of eligibility questions in the computer space that it is beginning to resemble an exclusive test. As a practical result, under current law, inventions not clearly intended to increase computer performance metrics are far more susceptible to invalidation under § 101. Many artificial intelligence patents, by contrast, are directed to new capabilities or qualitative improvements and are therefore in unwarranted jeopardy.

A. *Recent Cases Place an Increased Emphasis on Quantifiable Advances*

This focus on quantifiable computer improvements and discounting of qualitative improvements has accelerated since *Alice*, driven primarily by *Enfish v. Microsoft*.[73] The claims upheld in *Enfish* related to "an innovative logical model for a computer database" that explained "how the various elements of information are related to one another."[74] By using a "self-referential model" that "can store all entity types in a single table" and "can define the table's columns by rows in that same table," a database using the claimed invention could store certain types of data more effectively and could be searched more quickly.[75] The Federal Circuit distinguished the invention from the methods invalidated in *Alice* by describing it as "a specific improvement to the way computers operate," "an improvement in the functioning of a computer," and "a specific implementation of a solution to a problem in the software arts."[76] The Federal Circuit therefore found the challenged claims patent-eligible at *Alice* step one, although it noted that the concept of a specific improvement to computer functionality could potentially be relevant under *Alice* step two as well.[77] In particular, the Federal Circuit appears to

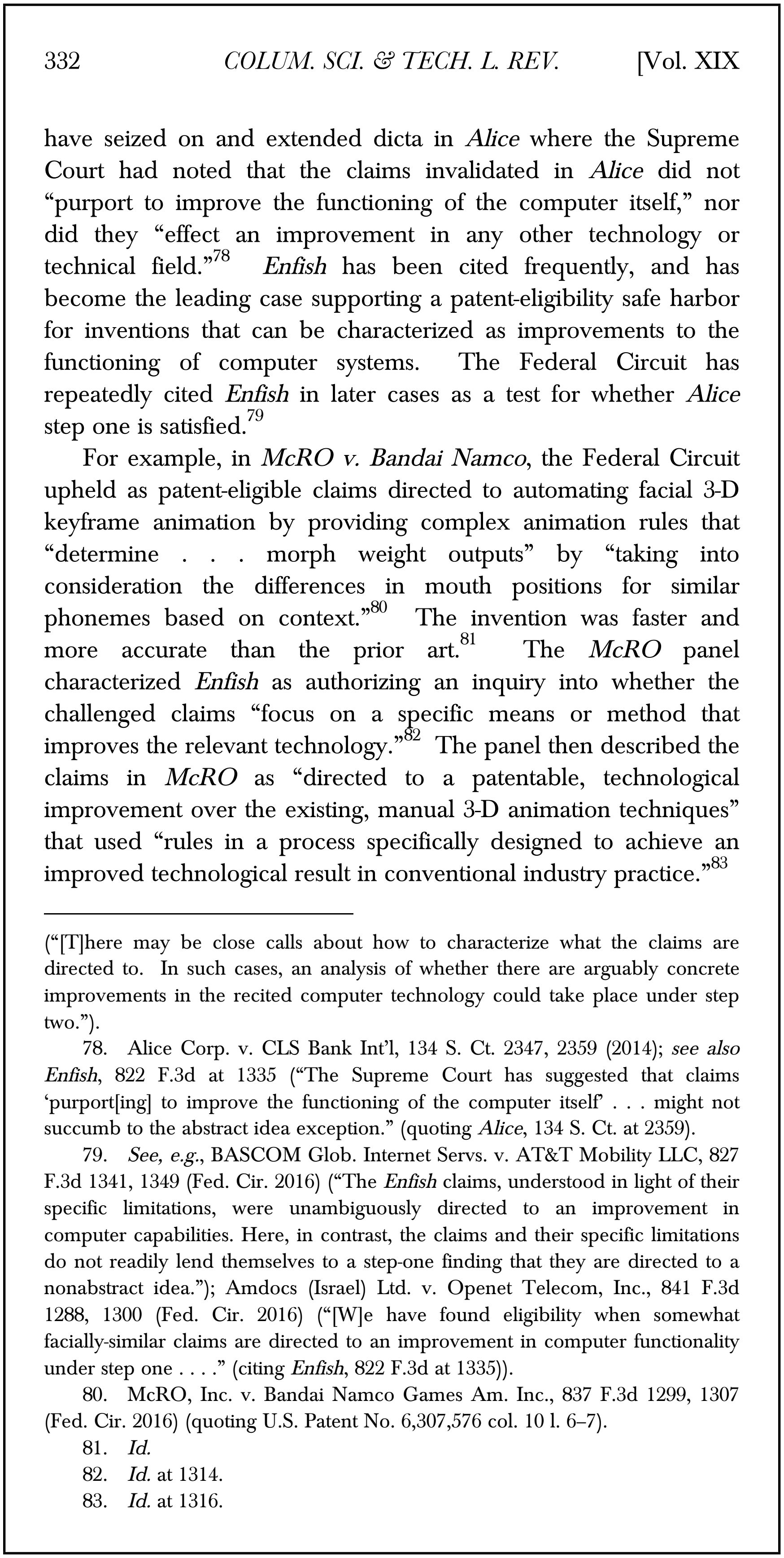
[73] Enfish, LLC v. Microsoft Corp., 822 F.3d 1327 (Fed. Cir. 2016).

[74] *Id.* at 1330.

[75] *Id.* at 1332–33.

[76] *Id.* at 1336–39.

[77] *Id.* at 1335 ("[W]e find it relevant to ask whether the claims are directed to an improvement to computer functionality versus being directed to an abstract idea, even at the first step of the *Alice* analysis."); *id. at* 1339



have seized on and extended dicta in *Alice* where the Supreme Court had noted that the claims invalidated in *Alice* did not “purport to improve the functioning of the computer itself,” nor did they “effect an improvement in any other technology or technical field.”[78] *Enfish* has been cited frequently, and has become the leading case supporting a patent-eligibility safe harbor for inventions that can be characterized as improvements to the functioning of computer systems. The Federal Circuit has repeatedly cited *Enfish* in later cases as a test for whether *Alice* step one is satisfied.[79]

For example, in *McRO v. Bandai Namco*, the Federal Circuit upheld as patent-eligible claims directed to automating facial 3-D keyframe animation by providing complex animation rules that “determine . . . morph weight outputs” by “taking into consideration the differences in mouth positions for similar phonemes based on context.”[80] The invention was faster and more accurate than the prior art.[81] The *McRO* panel characterized *Enfish* as authorizing an inquiry into whether the challenged claims “focus on a specific means or method that improves the relevant technology.”[82] The panel then described the claims in *McRO* as “directed to a patentable, technological improvement over the existing, manual 3-D animation techniques” that used “rules in a process specifically designed to achieve an improved technological result in conventional industry practice.”[83]

(“[T]here may be close calls about how to characterize what the claims are directed to. In such cases, an analysis of whether there are arguably concrete improvements in the recited computer technology could take place under step two.”).

[78] Alice Corp. v. CLS Bank Int’l, 134 S. Ct. 2347, 2359 (2014); *see also Enfish*, 822 F.3d at 1335 (“The Supreme Court has suggested that claims ‘purport[ing] to improve the functioning of the computer itself’ . . . might not succumb to the abstract idea exception.” (quoting *Alice*, 134 S. Ct. at 2359).

[79] *See, e.g.*, BASCOM Glob. Internet Servs. v. AT&T Mobility LLC, 827 F.3d 1341, 1349 (Fed. Cir. 2016) (“The *Enfish* claims, understood in light of their specific limitations, were unambiguously directed to an improvement in computer capabilities. Here, in contrast, the claims and their specific limitations do not readily lend themselves to a step-one finding that they are directed to a nonabstract idea.”); Amdocs (Israel) Ltd. v. Openet Telecom, Inc., 841 F.3d 1288, 1300 (Fed. Cir. 2016) (“[W]e have found eligibility when somewhat facially-similar claims are directed to an improvement in computer functionality under step one . . . .” (citing *Enfish*, 822 F.3d at 1335)).

[80] McRO, Inc. v. Bandai Namco Games Am. Inc., 837 F.3d 1299, 1307 (Fed. Cir. 2016) (quoting U.S. Patent No. 6,307,576 col. 10 l. 6–7).

[81] *Id.*

[82] *Id.* at 1314.

[83] *Id.* at 1316.

In *Thales v. United States*, the Federal Circuit continued its emphasis on whether the challenged claims represented a technological improvement.[84] Thales' patent was for "an inertial tracking system for tracking the motion of an object relative to a moving reference frame."[85] By "directly measur[ing] the gravitational field in the platform frame" the invention enabled "track[ing] the position and orientation of the object within the moving platform without input from a vehicle attitude reference system or calculating orientation or position of the moving platform itself."[86] Relying primarily on an analogy to the claimed rubber curing process in *Diamond v. Diehr*, the Federal Circuit upheld the patentability of the challenged claims.[87] But *Diehr* was now filtered through the lens of *Alice* and *Enfish*. The Federal Circuit characterized *Diehr*'s holding in terms of technological improvement: "In terms of the modern day *Alice* test, the *Diehr* claims were directed to an improvement in the rubber curing process, not a mathematical formula."[88]

Enfish was reaffirmed and extended in *Visual Memory v. NVIDIA*, in which the Federal Circuit found that claims directed to an "improved computer memory system" were patent-eligible.[89] The Federal Circuit even phrased the test for *Alice* step one as the question from *Enfish* of "whether the claims are directed to an improvement to computer functionality versus being directed to an abstract idea."[90] In finding that the memory storage claims at issue were valid, the Federal Circuit noted their linkage to computer architecture rather than "the abstract idea of categorical data storage."[91] The Federal Circuit also highlighted the ability of the claimed invention to improve general processor performance.[92]

[84] Thales Visionix Inc. v. United States, 850 F.3d 1343 (Fed. Cir. 2017).

[85] *Id.* at 1344.

[86] *Id.* at 1345.

[87] *Id.* at 1348 ("For the purpose of evaluating patent eligibility, the '159 patent claims are nearly indistinguishable from the claims at issue in *Diehr*.").

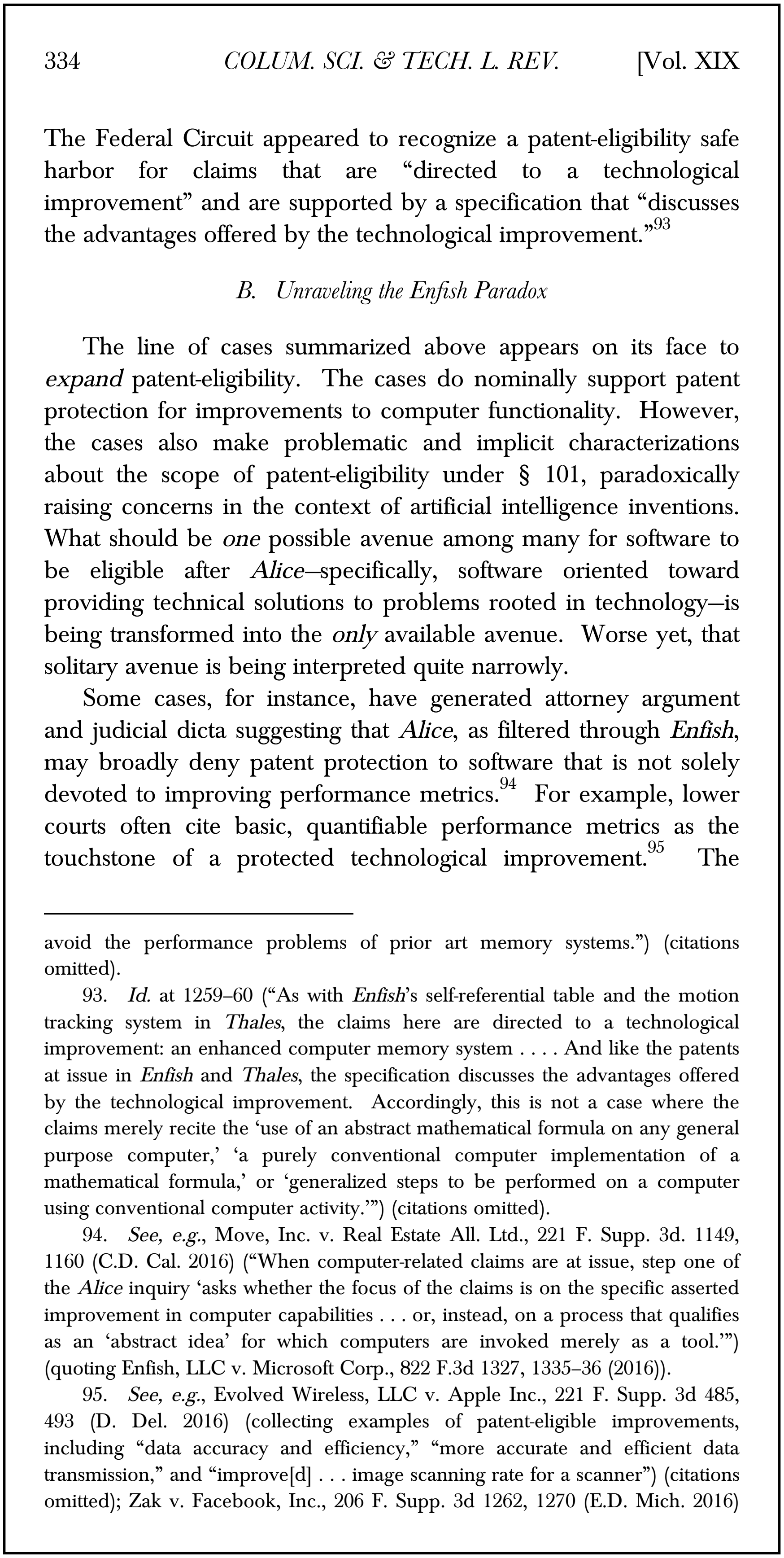
[88] *Id.*

[89] Visual Memory LLC v. NVIDIA Corp., 867 F.3d 1253, 1259 (Fed. Cir. 2017).

[90] *Id.* at 1258 (quoting Enfish, LLC v. Microsoft Corp., 822 F.3d 1327, 1335 (2016)).

[91] *Id.* at 1259 (citing, among other things, the claims' limitations to "programmable operational characteristics" and "storing certain types of data").

[92] *Id.* ("Although prior art memory systems possessed the flexibility to operate with multiple different processors, this one-size-fits-all approach frequently caused a tradeoff in processor performance. The '740 patent's teachings obviate the need to design a separate memory system for each type of processor, which proved to be costly and inefficient, and, at the same time,



The Federal Circuit appeared to recognize a patent-eligibility safe harbor for claims that are "directed to a technological improvement" and are supported by a specification that "discusses the advantages offered by the technological improvement."[93]

B. Unraveling the Enfish Paradox

The line of cases summarized above appears on its face to *expand* patent-eligibility. The cases do nominally support patent protection for improvements to computer functionality. However, the cases also make problematic and implicit characterizations about the scope of patent-eligibility under § 101, paradoxically raising concerns in the context of artificial intelligence inventions. What should be *one* possible avenue among many for software to be eligible after *Alice*—specifically, software oriented toward providing technical solutions to problems rooted in technology—is being transformed into the *only* available avenue. Worse yet, that solitary avenue is being interpreted quite narrowly.

Some cases, for instance, have generated attorney argument and judicial dicta suggesting that *Alice*, as filtered through *Enfish*, may broadly deny patent protection to software that is not solely devoted to improving performance metrics.[94] For example, lower courts often cite basic, quantifiable performance metrics as the touchstone of a protected technological improvement.[95] The

avoid the performance problems of prior art memory systems.") (citations omitted).

[93] *Id.* at 1259–60 ("As with *Enfish*'s self-referential table and the motion tracking system in *Thales*, the claims here are directed to a technological improvement: an enhanced computer memory system . . . . And like the patents at issue in *Enfish* and *Thales*, the specification discusses the advantages offered by the technological improvement. Accordingly, this is not a case where the claims merely recite the 'use of an abstract mathematical formula on any general purpose computer,' 'a purely conventional computer implementation of a mathematical formula,' or 'generalized steps to be performed on a computer using conventional computer activity.'") (citations omitted).

[94] *See, e.g.*, Move, Inc. v. Real Estate All. Ltd., 221 F. Supp. 3d. 1149, 1160 (C.D. Cal. 2016) ("When computer-related claims are at issue, step one of the *Alice* inquiry 'asks whether the focus of the claims is on the specific asserted improvement in computer capabilities . . . or, instead, on a process that qualifies as an 'abstract idea' for which computers are invoked merely as a tool.'") (quoting Enfish, LLC v. Microsoft Corp., 822 F.3d 1327, 1335–36 (2016)).

[95] *See, e.g.*, Evolved Wireless, LLC v. Apple Inc., 221 F. Supp. 3d 485, 493 (D. Del. 2016) (collecting examples of patent-eligible improvements, including "data accuracy and efficiency," "more accurate and efficient data transmission," and "improve[d] . . . image scanning rate for a scanner") (citations omitted); Zak v. Facebook, Inc., 206 F. Supp. 3d 1262, 1270 (E.D. Mich. 2016)

Federal Circuit has repeated in later cases a distinction, from *Enfish* itself, between protected improvements and unprotected computer-implemented inventions for "economic or other tasks for which a computer is used in its ordinary capacity."[96] A district court contrasted patent-eligible "solutions to computer-centric problems" with "performing abstract ideas in a digital medium."[97]

Lower courts have characterized this as deciding in *Alice* step one whether the challenged claims "are directed to an abstract idea or a specific improvement in computer capabilities"—i.e., that the opposite of an abstract idea is only a technological improvement and nothing else.[98] However, if read in this way, patent protection for artificial intelligence could largely evaporate. Artificial intelligence, after all, is not primarily concerned with making computers better at tasks that they already do. The quantitative benchmarks available in other applications of computer technology are unlikely to be available in claims directed to qualitative improvements in computer functionality—inventions that expand upon the ability of a computer to "see," or to "hear," or even to render informed judgments with incomplete information about subjective subject matter.

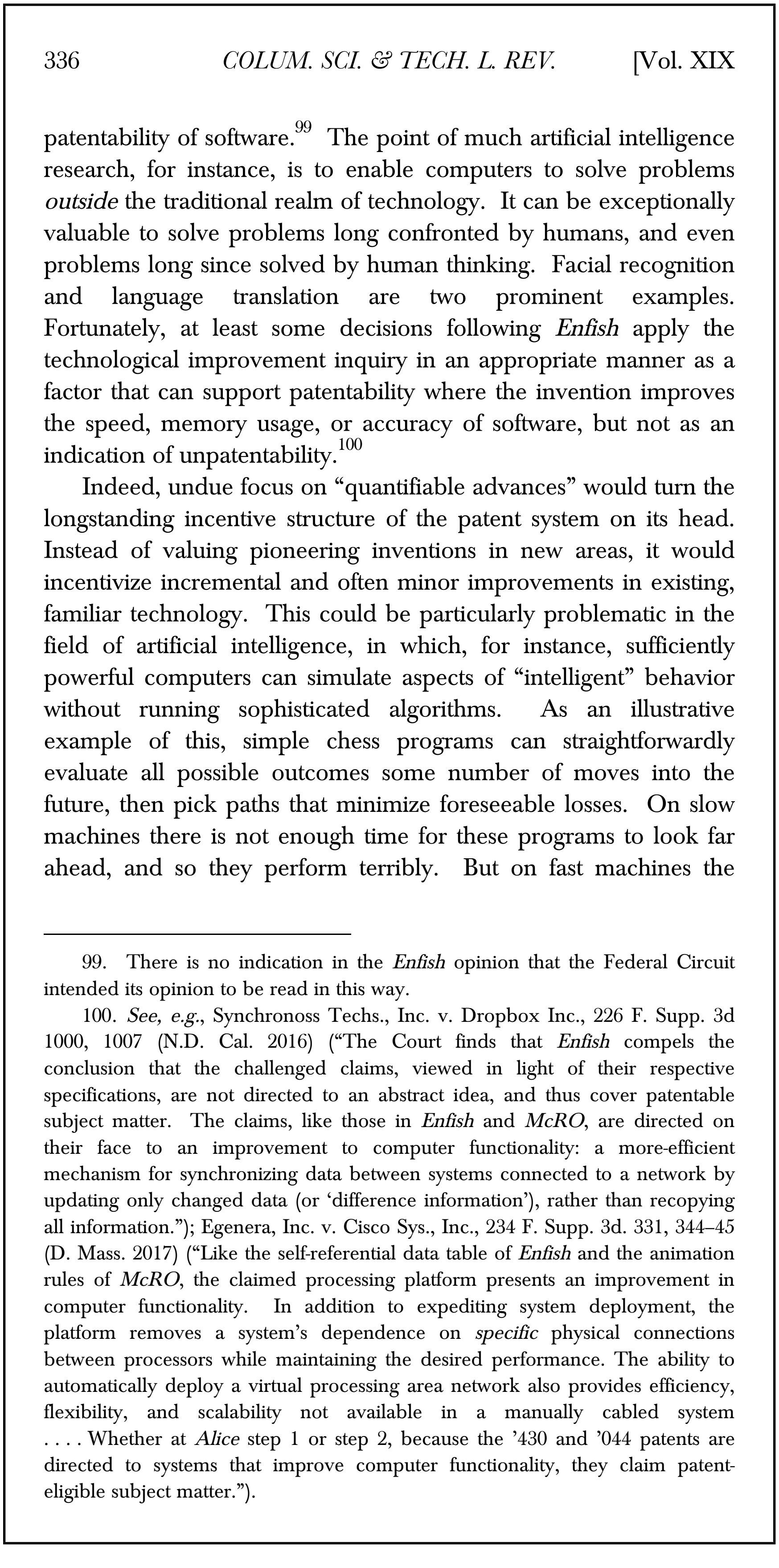
This Article submits that this is not a proper reading of the *Enfish* line of cases, which do not provide an exclusive test for the

("Unlike the claims in *Enfish*, the claim in the present case, in consideration of its limitations, does not unambiguously purport to increase speed, improve storage, or improve functionality of the computer itself.").

[96] *See, e.g.*, Secured Mail Sols. LLC v. Universal Wilde, Inc., 873 F.3d 905, 910 (Fed. Cir. 2017) (citing Enfish, LLC v. Microsoft Corp., 822 F.3d 1327, 1336) ("The court in *Enfish* held the claims relating to a computer database implementation to be patent-eligible under *Alice* step one because the claims focused on an improvement to computer functionality itself, *not on economic or other tasks for which a computer is used in its ordinary capacity*.") (emphasis added)).

[97] Virginia Innovation Scis, Inc. v. Amazon.com, Inc., 227 F. Supp. 3d 582, 597 (E.D. Va. 2017) ("Comparing this to the other recent cases, while *McRO* and *Enfish* are efforts to improve data processing, the patents in *TLI* and this case are only possible *because* of data processing. Therefore, they are akin to performing abstract ideas in a digital medium rather than creating solutions to computer-centric problems.").

[98] *Evolved Wireless*, 221 F. Supp. 3d at 492 (D. Del. 2016) ("Accordingly, the court will consider under the first step of *Alice* whether the '916 and '481 patents are directed to an abstract idea or a specific improvement in computer capabilities . . . . In determining whether the mathematical algorithms disclosed in the patents at issue are directed to an abstract idea or technological improvement, the court finds instructive cases addressing similar technological problems and solutions.").



patentability of software.[99] The point of much artificial intelligence research, for instance, is to enable computers to solve problems *outside* the traditional realm of technology. It can be exceptionally valuable to solve problems long confronted by humans, and even problems long since solved by human thinking. Facial recognition and language translation are two prominent examples. Fortunately, at least some decisions following *Enfish* apply the technological improvement inquiry in an appropriate manner as a factor that can support patentability where the invention improves the speed, memory usage, or accuracy of software, but not as an indication of unpatentability.[100]

Indeed, undue focus on "quantifiable advances" would turn the longstanding incentive structure of the patent system on its head. Instead of valuing pioneering inventions in new areas, it would incentivize incremental and often minor improvements in existing, familiar technology. This could be particularly problematic in the field of artificial intelligence, in which, for instance, sufficiently powerful computers can simulate aspects of "intelligent" behavior without running sophisticated algorithms. As an illustrative example of this, simple chess programs can straightforwardly evaluate all possible outcomes some number of moves into the future, then pick paths that minimize foreseeable losses. On slow machines there is not enough time for these programs to look far ahead, and so they perform terribly. But on fast machines the

[99] There is no indication in the *Enfish* opinion that the Federal Circuit intended its opinion to be read in this way.

[100] *See, e.g.*, Synchronoss Techs., Inc. v. Dropbox Inc., 226 F. Supp. 3d 1000, 1007 (N.D. Cal. 2016) ("The Court finds that *Enfish* compels the conclusion that the challenged claims, viewed in light of their respective specifications, are not directed to an abstract idea, and thus cover patentable subject matter. The claims, like those in *Enfish* and *McRO*, are directed on their face to an improvement to computer functionality: a more-efficient mechanism for synchronizing data between systems connected to a network by updating only changed data (or 'difference information'), rather than recopying all information."); Egenera, Inc. v. Cisco Sys., Inc., 234 F. Supp. 3d. 331, 344–45 (D. Mass. 2017) ("Like the self-referential data table of *Enfish* and the animation rules of *McRO*, the claimed processing platform presents an improvement in computer functionality. In addition to expediting system deployment, the platform removes a system's dependence on *specific* physical connections between processors while maintaining the desired performance. The ability to automatically deploy a virtual processing area network also provides efficiency, flexibility, and scalability not available in a manually cabled system . . . . Whether at *Alice* step 1 or step 2, because the '430 and '044 patents are directed to systems that improve computer functionality, they claim patent-eligible subject matter.").

same programs can defeat skilled opponents.[101] Under what is currently the prevailing reading of *Enfish*, the quantifiable performance improvement obtained by simply running the program on the faster machine would weigh in favor of patentability, even though this is arguably not innovation of the sort our patent system should be geared to reward.[102] Unwarranted industry focus on improving routine performance metrics, such as would be encouraged by treating the *Enfish* safe harbor as an exclusive test, would likely discourage fundamental innovation in artificial intelligence.

There is also good reason to treat *Enfish*'s reference to "specific improvements in computer capabilities" broadly in a manner that does not require quantitative advances. Even *Enfish* implied that a protected technological improvement may be to "logical structures and processes" of software, which does not necessarily exclude qualitative improvements.[103] And *Enfish* recognized that "in other cases involving computer-related claims, there may be close calls about how to characterize what the claims are directed to."[104] *Enfish* further noted in passing that one improvement offered by the challenged claims was "increased flexibility" in a database system.[105] The Federal Circuit in *BASCOM v. AT&T* similarly observed, when reading *Enfish* in light of *Alice*, that "it might become clear that the specific improvements in the recited computer technology go beyond well-understood, routine,

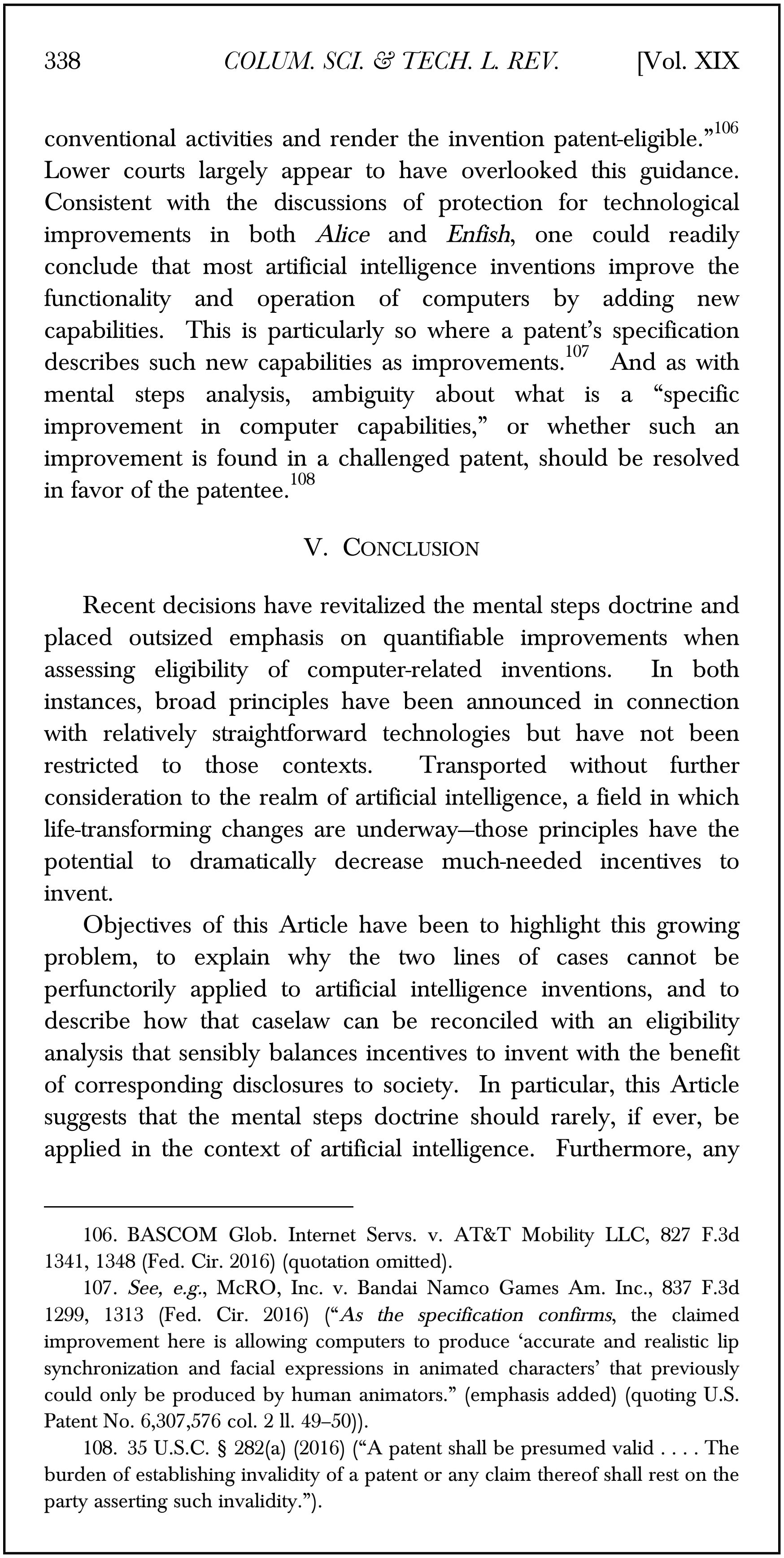
[101] On an idealized machine with unlimited processing power, the simple program would play perfect chess. Conversely, even sophisticated programs on machines with limited processing capability may be susceptible to "anti-computer" styles of play that would not fool skilled human opponents. *See, e.g.*, Tim Krabbé, *Defending Humanity's Honor* (2001), DE WEBSITE VAN TIM KRABBÉ, http://timkr.home.xs4all.nl/chess2/honor.htm (profiling a specialist in computer chess who "consistently beats the world's strongest commercial chess programs with a unique anti-computer style").

[102] 2015–2016 STUDY PANEL OF THE ONE HUNDRED YEAR STUDY ON ARTIFICIAL INTELLIGENCE, STANFORD UNIV., ARTIFICIAL INTELLIGENCE AND LIFE IN 2030 13 (Sept. 2016), https://ai100.stanford.edu/sites/default/files/ai100report10032016fnl_singles.pdf (citing contemporary criticism of IBM's Deep Blue chess computer that beat Gary Kasparov in 1997 as "a collection of 'brute force methods' that wasn't 'real intelligence'").

[103] Enfish, LLC v. Microsoft Corp., 822 F.3d 1327, 1339 (Fed. Cir. 2016) ("Much of the advancement made in computer technology consists of improvements to software that, by their very nature, may not be defined by particular physical features but rather by logical structures and processes. We do not see in *Bilski* or *Alice*, or our cases, an exclusion to patenting this large field of technological progress.").

[104] *Id.*

[105] *Id.* at 1337.



conventional activities and render the invention patent-eligible."[106] Lower courts largely appear to have overlooked this guidance. Consistent with the discussions of protection for technological improvements in both *Alice* and *Enfish*, one could readily conclude that most artificial intelligence inventions improve the functionality and operation of computers by adding new capabilities. This is particularly so where a patent's specification describes such new capabilities as improvements.[107] And as with mental steps analysis, ambiguity about what is a "specific improvement in computer capabilities," or whether such an improvement is found in a challenged patent, should be resolved in favor of the patentee.[108]

V. CONCLUSION

Recent decisions have revitalized the mental steps doctrine and placed outsized emphasis on quantifiable improvements when assessing eligibility of computer-related inventions. In both instances, broad principles have been announced in connection with relatively straightforward technologies but have not been restricted to those contexts. Transported without further consideration to the realm of artificial intelligence, a field in which life-transforming changes are underway—those principles have the potential to dramatically decrease much-needed incentives to invent.

Objectives of this Article have been to highlight this growing problem, to explain why the two lines of cases cannot be perfunctorily applied to artificial intelligence inventions, and to describe how that caselaw can be reconciled with an eligibility analysis that sensibly balances incentives to invent with the benefit of corresponding disclosures to society. In particular, this Article suggests that the mental steps doctrine should rarely, if ever, be applied in the context of artificial intelligence. Furthermore, any

[106] BASCOM Glob. Internet Servs. v. AT&T Mobility LLC, 827 F.3d 1341, 1348 (Fed. Cir. 2016) (quotation omitted).

[107] *See, e.g.*, McRO, Inc. v. Bandai Namco Games Am. Inc., 837 F.3d 1299, 1313 (Fed. Cir. 2016) ("*As the specification confirms*, the claimed improvement here is allowing computers to produce 'accurate and realistic lip synchronization and facial expressions in animated characters' that previously could only be produced by human animators." (emphasis added) (quoting U.S. Patent No. 6,307,576 col. 2 ll. 49–50)).

[108] 35 U.S.C. § 282(a) (2016) ("A patent shall be presumed valid . . . . The burden of establishing invalidity of a patent or any claim thereof shall rest on the party asserting such invalidity.").

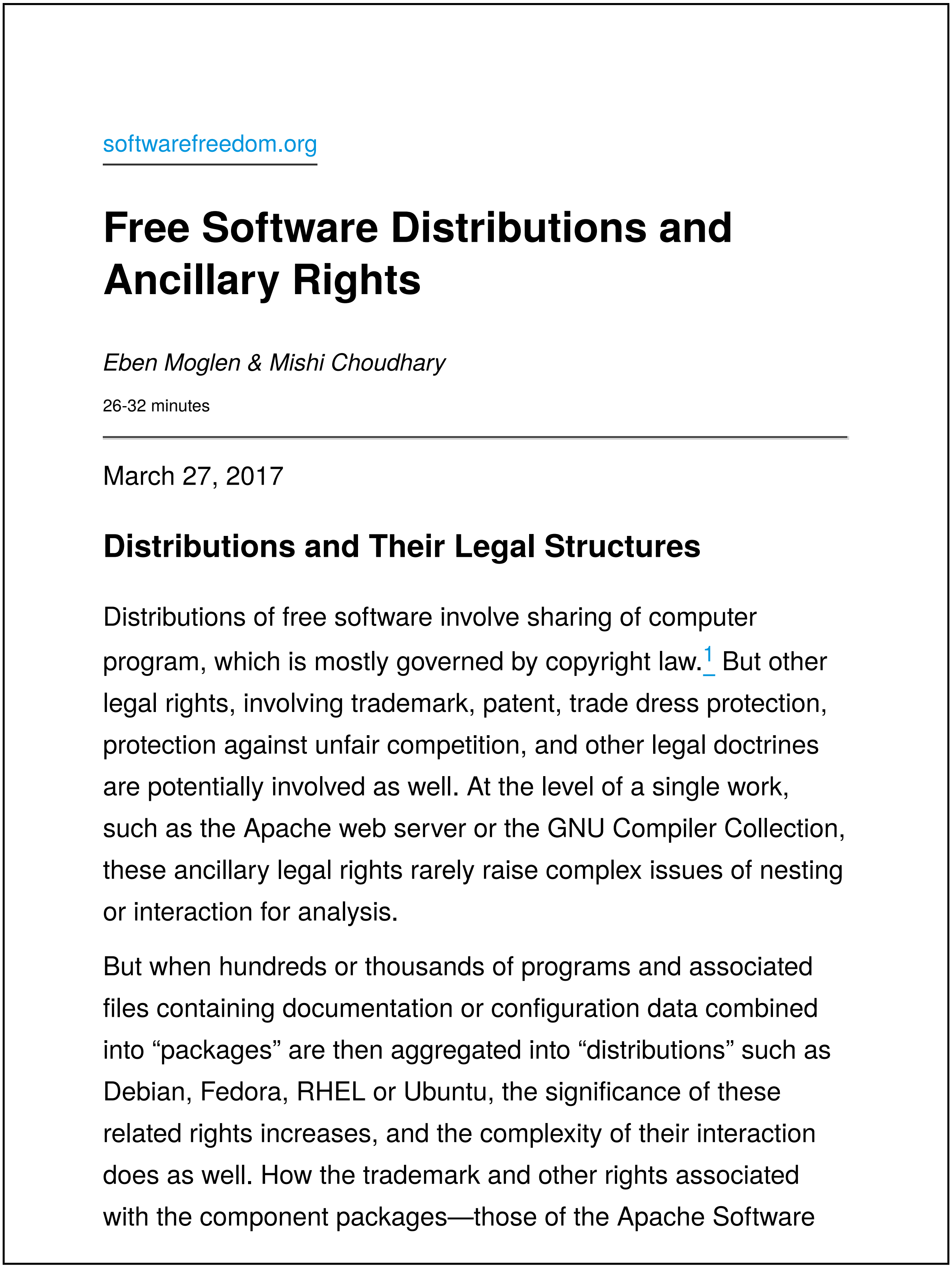
application of the doctrine should respect the fundamental difference between inventions that emulate or replace human thought and those that simply cover existing human activity. It also suggests that the emerging focus on "specific improvements in computer capabilities" should not transform what was intended as a non-exclusive safe harbor into the sole test for eligibility of computer-related inventions. A broad array of novel artificial intelligence techniques warrant protection irrespective of whether they provide readily calculable increases in conventional computer performance metrics. Applying these considerations when evaluating the patentability of inventions in the field of artificial intelligence, courts can once again provide an appropriate scope of protection for this important technology.

Part V

The End of One Era, the Start of Another

Chapter 25

Free Software Distributions and Ancillary Rights (Moglen and Choudhary)



softwarefreedom.org

Free Software Distributions and Ancillary Rights

Eben Moglen & Mishi Choudhary

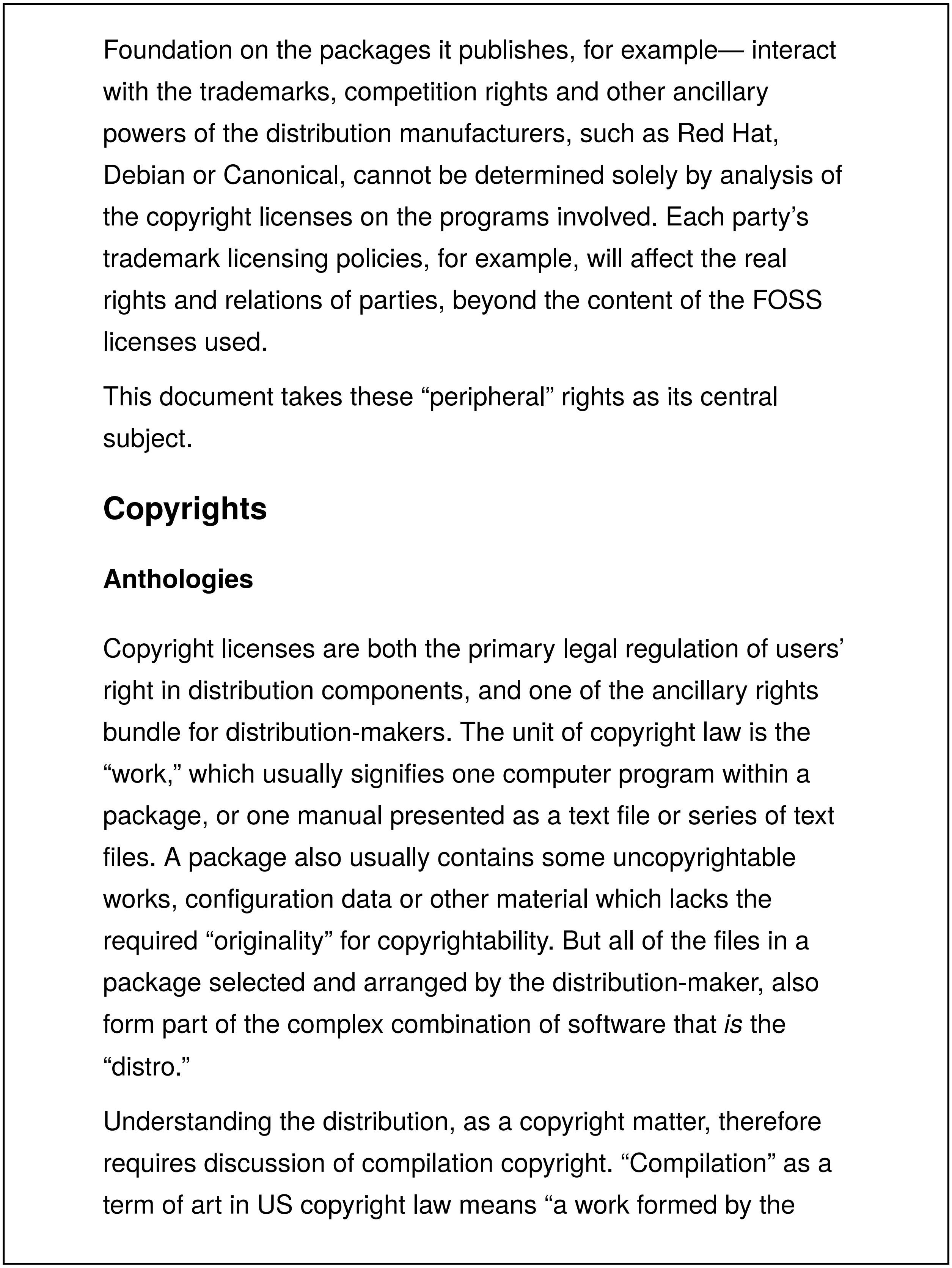
26-32 minutes

March 27, 2017

Distributions and Their Legal Structures

Distributions of free software involve sharing of computer program, which is mostly governed by copyright law.[1] But other legal rights, involving trademark, patent, trade dress protection, protection against unfair competition, and other legal doctrines are potentially involved as well. At the level of a single work, such as the Apache web server or the GNU Compiler Collection, these ancillary legal rights rarely raise complex issues of nesting or interaction for analysis.

But when hundreds or thousands of programs and associated files containing documentation or configuration data combined into "packages" are then aggregated into "distributions" such as Debian, Fedora, RHEL or Ubuntu, the significance of these related rights increases, and the complexity of their interaction does as well. How the trademark and other rights associated with the component packages—those of the Apache Software



Foundation on the packages it publishes, for example— interact with the trademarks, competition rights and other ancillary powers of the distribution manufacturers, such as Red Hat, Debian or Canonical, cannot be determined solely by analysis of the copyright licenses on the programs involved. Each party's trademark licensing policies, for example, will affect the real rights and relations of parties, beyond the content of the FOSS licenses used.

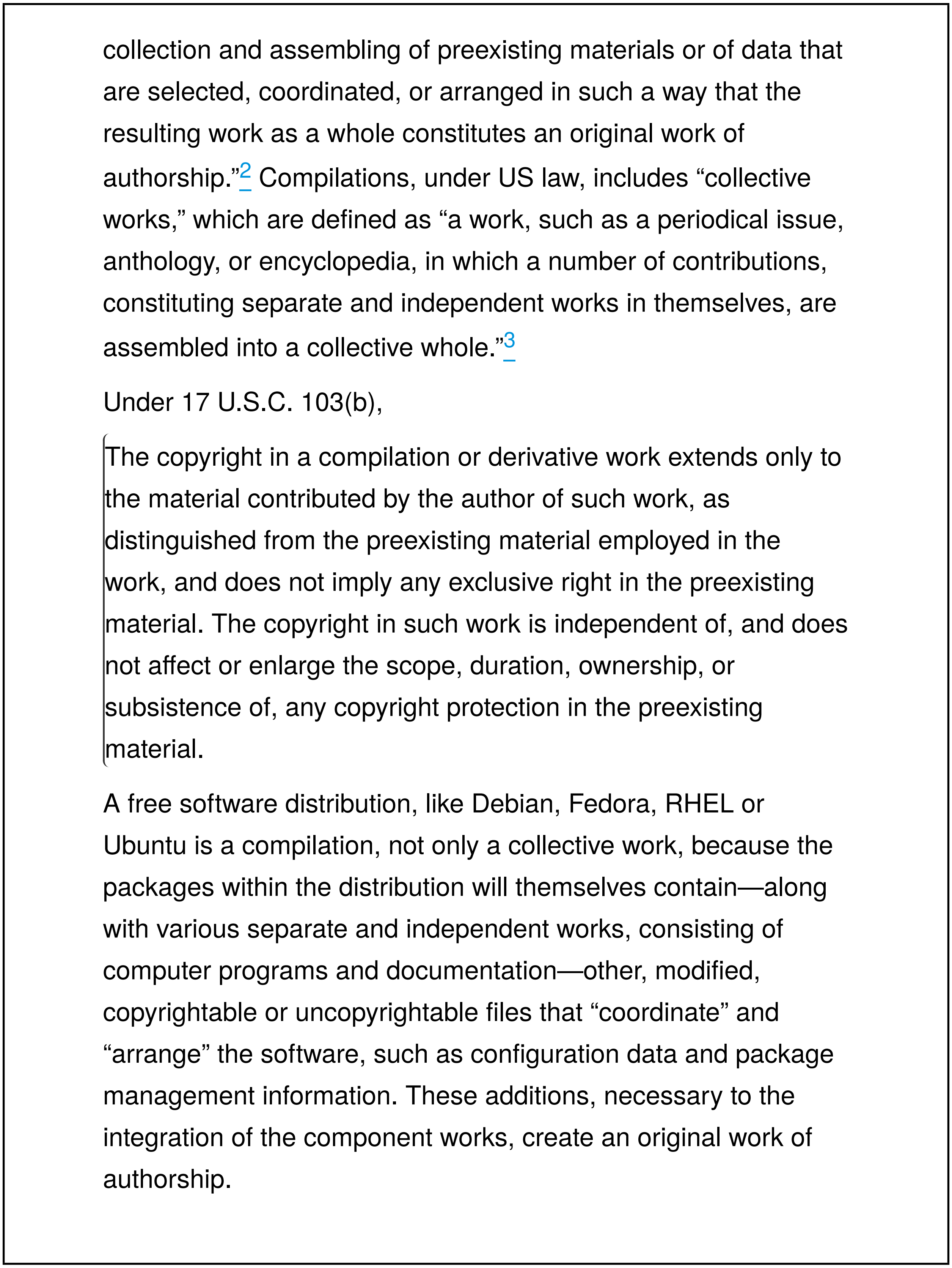
This document takes these "peripheral" rights as its central subject.

Copyrights

Anthologies

Copyright licenses are both the primary legal regulation of users' right in distribution components, and one of the ancillary rights bundle for distribution-makers. The unit of copyright law is the "work," which usually signifies one computer program within a package, or one manual presented as a text file or series of text files. A package also usually contains some uncopyrightable works, configuration data or other material which lacks the required "originality" for copyrightability. But all of the files in a package selected and arranged by the distribution-maker, also form part of the complex combination of software that *is* the "distro."

Understanding the distribution, as a copyright matter, therefore requires discussion of compilation copyright. "Compilation" as a term of art in US copyright law means "a work formed by the



collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship."[2] Compilations, under US law, includes "collective works," which are defined as "a work, such as a periodical issue, anthology, or encyclopedia, in which a number of contributions, constituting separate and independent works in themselves, are assembled into a collective whole."[3]

Under 17 U.S.C. 103(b),

> The copyright in a compilation or derivative work extends only to the material contributed by the author of such work, as distinguished from the preexisting material employed in the work, and does not imply any exclusive right in the preexisting material. The copyright in such work is independent of, and does not affect or enlarge the scope, duration, ownership, or subsistence of, any copyright protection in the preexisting material.

A free software distribution, like Debian, Fedora, RHEL or Ubuntu is a compilation, not only a collective work, because the packages within the distribution will themselves contain—along with various separate and independent works, consisting of computer programs and documentation—other, modified, copyrightable or uncopyrightable files that "coordinate" and "arrange" the software, such as configuration data and package management information. These additions, necessary to the integration of the component works, create an original work of authorship.

Copyright in a compilation gives the compiler exclusive rights to the selection and arrangement of the individual components, regardless of the copyright status of those components, and who holds copyright on them.[4] The compilation copyright neither ousts nor modifies the respective authors' exclusive rights, which must be licensed to the compilation-maker in order to create the compilation.

Selection and Arrangement

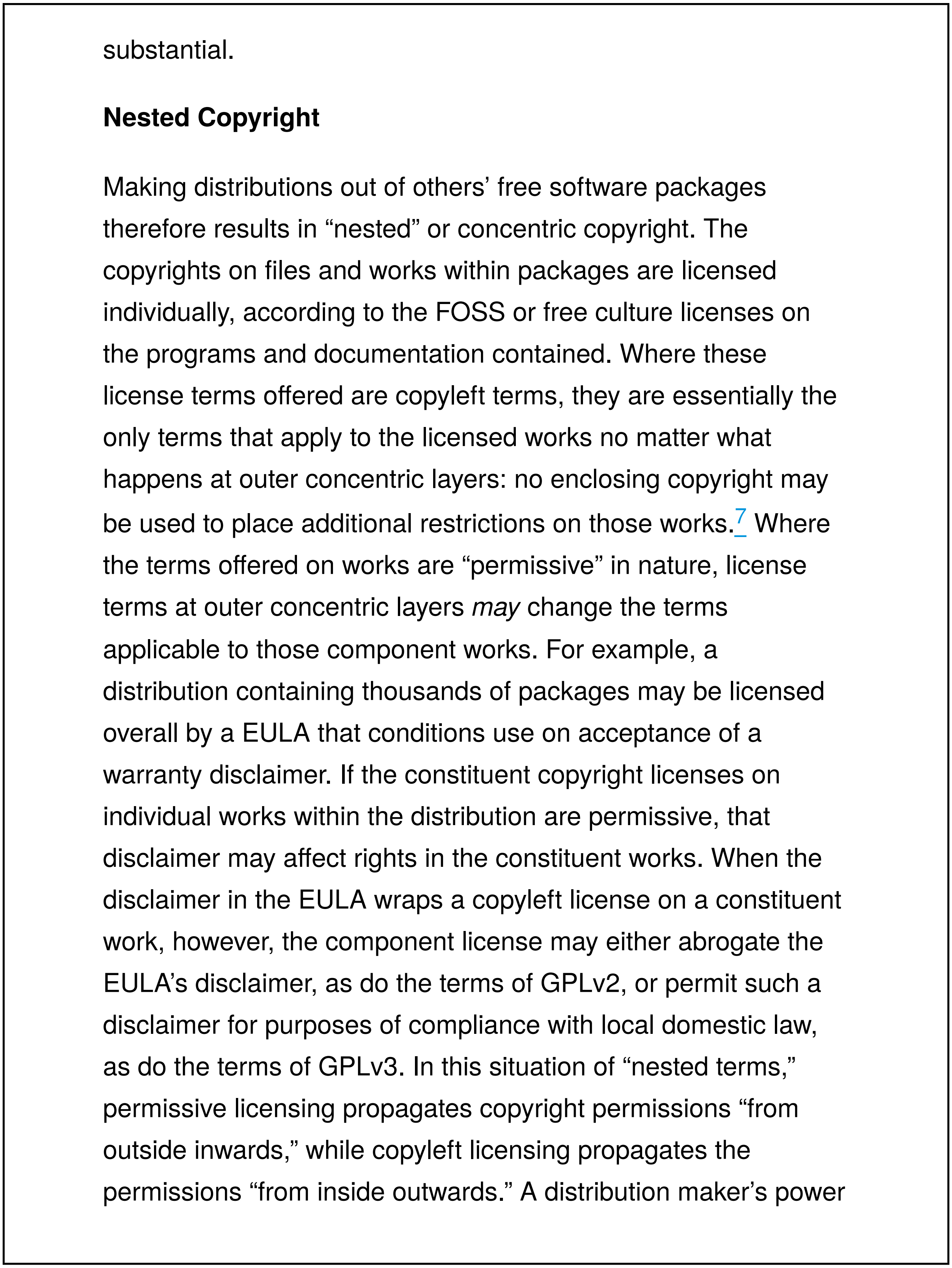
The activity of putting together a distribution of free software packages involves selecting and arranging packages; creating necessary intra-package and distribution-wide management and configuration information; compiling, processing or otherwise creating executable files from source code trees; and arranging the resulting distribution files for convenient acquisition, installation and updating by users. Some of these actions give rise to copyrightability in the resulting aggregation of software, in the same way that the choice and sequencing of literary works in an anthology gives rise to copyright.[5] The decision to include the Apache web server, or the lighttpd web server, or both, in a distribution is part of the selection and arrangement that gives rise to the "anthology copyright" on the distro as a whole. Other portions of the distribution-maker's activity result in purely functional files or mechanical transformations of others' copyrighted works, valuable to the user but not subject to copyright because not an "original work of authorship," such as the binaries compiled from others' source code. Some results from these latter activities may give rise to other ancillary rights, under, for example, domestic law of legal protection for databases created and maintained as part of the distribution.[6]

In-House Code

The process of making distributions involves no small amount of additional programming. Maintenance to package code, works specific to the distribution, activities and projects specifically produced or sponsored by the distribution-maker—all comprise a significant fraction of the software in major distributions. These activities produce their own independent copyrights in many instances. These work- or package-level copyrights form another part of the web of ancillary rights held by distribution makers.

Configuration copyright

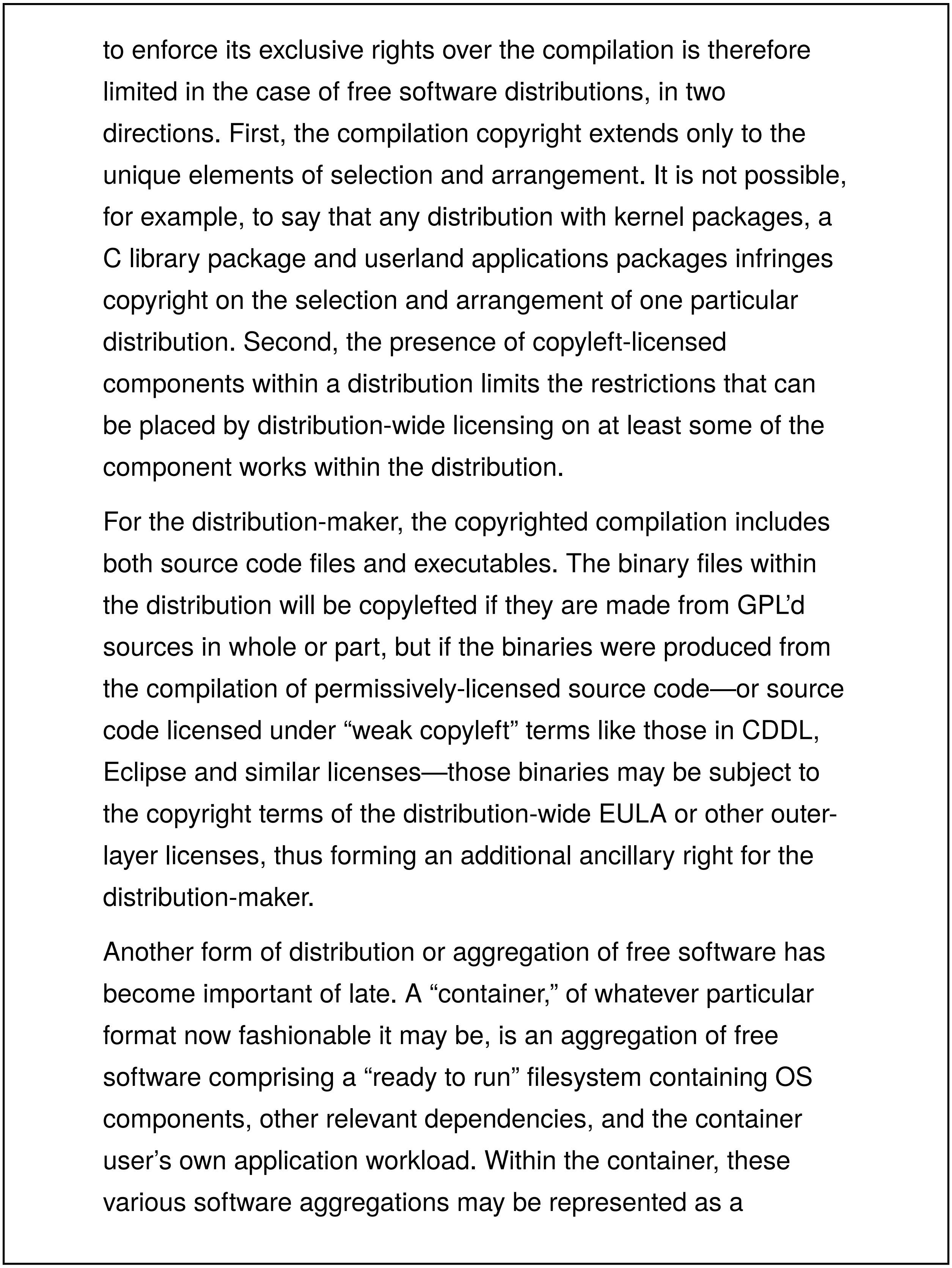
Most packages in a distribution contain configuration data, specifying details of the package's operation when installed. Default options, file pathnames, device- or system-specific information and other functionally useful details are contained in these configuration files, which are often annotated by text explaining for human readers what the configuration values mean and how to modify them. These configuration files are thinly copyrightable, if they can be said to fall within copyright scope at all. The sum of these configuration files—which allow all the packages in a distribution to work together smoothly because they are compatibly configured—expresses precisely the creative selection and arrangement of the distro. So while the copyrightability of the individual files may be doubtful or scant, the overall effect of these fragmented copyrights is



substantial.

Nested Copyright

Making distributions out of others' free software packages therefore results in "nested" or concentric copyright. The copyrights on files and works within packages are licensed individually, according to the FOSS or free culture licenses on the programs and documentation contained. Where these license terms offered are copyleft terms, they are essentially the only terms that apply to the licensed works no matter what happens at outer concentric layers: no enclosing copyright may be used to place additional restrictions on those works.[7] Where the terms offered on works are "permissive" in nature, license terms at outer concentric layers *may* change the terms applicable to those component works. For example, a distribution containing thousands of packages may be licensed overall by a EULA that conditions use on acceptance of a warranty disclaimer. If the constituent copyright licenses on individual works within the distribution are permissive, that disclaimer may affect rights in the constituent works. When the disclaimer in the EULA wraps a copyleft license on a constituent work, however, the component license may either abrogate the EULA's disclaimer, as do the terms of GPLv2, or permit such a disclaimer for purposes of compliance with local domestic law, as do the terms of GPLv3. In this situation of "nested terms," permissive licensing propagates copyright permissions "from outside inwards," while copyleft licensing propagates the permissions "from inside outwards." A distribution maker's power



to enforce its exclusive rights over the compilation is therefore limited in the case of free software distributions, in two directions. First, the compilation copyright extends only to the unique elements of selection and arrangement. It is not possible, for example, to say that any distribution with kernel packages, a C library package and userland applications packages infringes copyright on the selection and arrangement of one particular distribution. Second, the presence of copyleft-licensed components within a distribution limits the restrictions that can be placed by distribution-wide licensing on at least some of the component works within the distribution.

For the distribution-maker, the copyrighted compilation includes both source code files and executables. The binary files within the distribution will be copylefted if they are made from GPL'd sources in whole or part, but if the binaries were produced from the compilation of permissively-licensed source code—or source code licensed under “weak copyleft” terms like those in CDDL, Eclipse and similar licenses—those binaries may be subject to the copyright terms of the distribution-wide EULA or other outer-layer licenses, thus forming an additional ancillary right for the distribution-maker.

Another form of distribution or aggregation of free software has become important of late. A “container,” of whatever particular format now fashionable it may be, is an aggregation of free software comprising a “ready to run” filesystem containing OS components, other relevant dependencies, and the container user’s own application workload. Within the container, these various software aggregations may be represented as a collection of compressed archive files, for example, all of which, when expanded, present the full filesystem intended for execution. The container can be compared roughly to a “live CD” filesystem, or a bootable USB filesystem, which have been and remain popular modes of aggregate free software distribution.

In a container, as in one of these conceptual predecessors, the copyleft binary components are—from the point of view of GPLv2, for example—placed in “mere aggregation” with other non-GPL binaries, not in a combination triggering copyleft. So being in the same container with a GPL’d binary doesn’t propagate the copyleft “outward.” (The same requirements to provide or offer source code for the GPL’d executables apply, however, if A provides a “container” to B as would be imposed if A provided a CD or USB key to B with the same GPL’d executables on it.)

But the selection and arrangement of works aggregated in a container could, even if the container was made under program control, give rise to “anthology” copyright over the aggregate. The terms on which this anthology is licensed could theoretically modify the terms on permissively-licensed works within the container, as we have seen in other contexts.

So far as we are aware, no container-related anthology-level copyright issues have so far occurred in practice. Parties have been more concerned with a non-existent copyleft inheritance issue, where GPL’d binaries are included in containers also containing proprietary or other non-copyleft executables. Once this uncertainty has subsided, however, the issue of EULA-like enclosing licenses on containers is likely to arise in one or more commercial contexts.

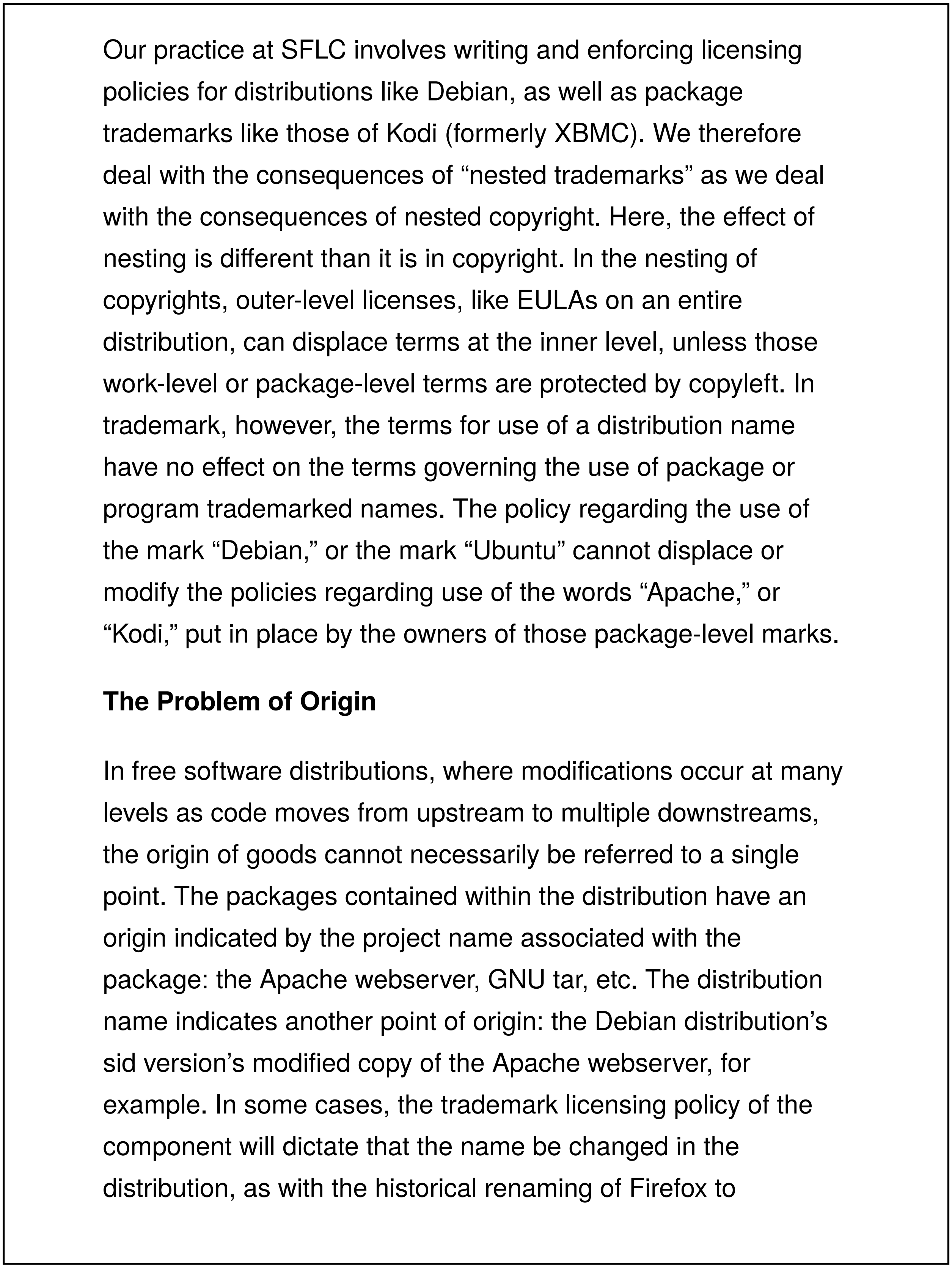
Trademarks

The essence of trademark is the association of a name with the source or origin of goods in the marketplace.[8] The property right in the name given to the trademark holder is balanced by the social advantage consumers derive from the assurance of quality names convey. Accordingly, in the United States, the essence of trademark liability is the creation of "confusion" by disturbance of the relationship between the known mark and the inferred source, origin, and quality of goods.[9]

Descriptive names can be trademarks, in the US, only to the extent that the descriptive name has acquired "secondary meaning," that is, has come to be associated with the particular source or origin of goods through use and acceptance in the marketplace. Even an arbitrary name, such as Kleenex, can become "genericized" through the process of broad social use, such that the secondary meaning associated with a particular producer of paper in which to blow your nose has evaporated over time. Unlike copyrights, trademark rights are subject to defeasement through failure to protect the mark. The trademark owner's failure to prevent confusion can lead to loss of rights.[10] Trademark owners are therefore required to take measures to "police" their marks, which is why trademark enforcement often occurs in situations that seem in isolation oppressive or trivial.

Free software programs and packages are named; the names chosen are often registered as trademarks.[11] In order to avoid possible loss of the mark through a finding of “naked licensing,” the projects using trademarked names need to make licensing policies, stating clearly the terms on which the marks, whether textual or graphical, may be applied to modified or redistributed versions. They need to be prepared to show that their policies are enforced.[12] Unlike commercial licensors, who belligerantly enforce their marks, it is often in the interest of free software projects to make liberal trademark policies, allowing those downstream who modify or repackage their code to make use of their marks. The origin of the free software is explicitly designated, by the use of the package’s trademarked name, while the fact that the distribution and some or all component packages have been modified is also indicated in the modified version’s documentation and source code. Clearly indicating the origin of modifications while retaining the name of the package obviates risk of confusion.

Distributions, in turn, also have names, which are very likely to be registered as trademarks. They too have licensing policies, which govern the conditions under which the names can be used to identify, among other things, compilations of software that originate from the trademarked distribution, but which may have been modified downstream. Community distributions, such as Debian, are likely to have different trademark licensing policies than commercial distributions. The interests of the two kinds of parties dictate different balances between encouraging modification and redistribution, on the one hand, and protecting asset values, including of intangible property rights in names, on the other.



Our practice at SFLC involves writing and enforcing licensing policies for distributions like Debian, as well as package trademarks like those of Kodi (formerly XBMC). We therefore deal with the consequences of “nested trademarks” as we deal with the consequences of nested copyright. Here, the effect of nesting is different than it is in copyright. In the nesting of copyrights, outer-level licenses, like EULAs on an entire distribution, can displace terms at the inner level, unless those work-level or package-level terms are protected by copyleft. In trademark, however, the terms for use of a distribution name have no effect on the terms governing the use of package or program trademarked names. The policy regarding the use of the mark “Debian,” or the mark “Ubuntu” cannot displace or modify the policies regarding use of the words “Apache,” or “Kodi,” put in place by the owners of those package-level marks.

The Problem of Origin

In free software distributions, where modifications occur at many levels as code moves from upstream to multiple downstreams, the origin of goods cannot necessarily be referred to a single point. The packages contained within the distribution have an origin indicated by the project name associated with the package: the Apache webserver, GNU tar, etc. The distribution name indicates another point of origin: the Debian distribution’s sid version’s modified copy of the Apache webserver, for example. In some cases, the trademark licensing policy of the component will dictate that the name be changed in the distribution, as with the historical renaming of Firefox to

Iceweasel in Debian.[13]

At each level of enclosure, from the single program through the package to the distribution, the interest of the trademark holder is in preventing uses of the name which may confuse the ultimate user about the origin of the software, both in order to protect the user's quality perception of the "brand" established by the name, and to prevent harmful associations. The latter motive is important, for example, when a media player like Kodi is distributed by makers of "distributions" who add plug-ins to the player that facilitate the acquisition of copyright-infringing media files.

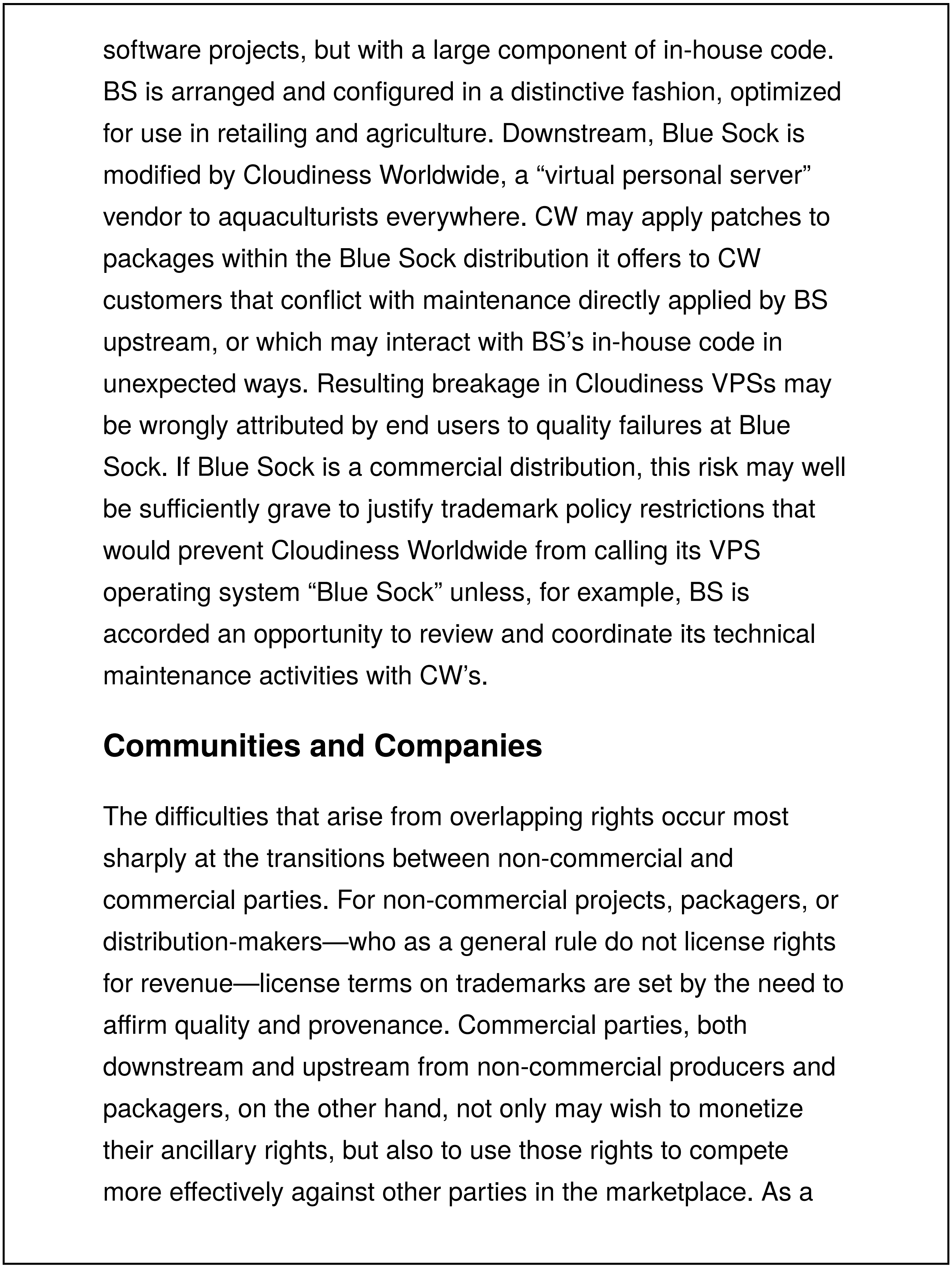
For this reason, the trademark licensing policies of both package and distribution markholders may require that materials bearing trademarks (graphical elements in UI design, documentation, etc., for example) must be removed from modified and redistributed versions, to prevent "confusion" as to origin. If such requirements are imposed by policy, but made difficult to implement, they could function as barriers to the redistribution of the free software itself, which—in the case of copylefted software—would imply a prohibited imposition of additional restrictions. In such cases, copyleft requires that the materials added midstream by the trademark owner to preexisting copylefted software be so segregated that they can with reasonable effort be removed by downstream modifiers. Nothing, on the other hand, ever *requires* a downstream party to retain a trademarked name it would rather remove.

Confusion Prevention—Package Quality Control

The rules that determine what names can be used for modified versions of FOSS programs are found not in the copyright licenses, but in the trademark licensing policy of the entity that controls each mark. The licensor's primary incentive is to prevent the quality assumptions that have been built into the mark from being deprecated downstream. This incentive prevents projects and packagers from being as permissive in trademark licensing as they are in their copyright licenses. It is sufficient, in most free software copyright licenses, that modified versions be so designated [14] But if program FOO, whose name is trademarked by its authors, is modified downstream and the modifications are properly designated in the source code as required by FOO's license, that does not put at ease the minds of FOO's developers if the modifications contain serious bugs that will cause users of modified FOO to experience harm, thus deprecating FOO's reputation for quality. For this reason, the FOO project's trademark policy could require that modification patchsets be reviewed by FOO's developers before they will license use of the trademark FOO as the name of the modified version.

Confusion Prevention—Distribution Conflict

Trademark licensing policies can also play a role in avoiding confusion not only in end-user perception but also in technical coordination at the distribution level. Let's take an imaginary trademarked free software distribution called Blue Sock, containing the Linux kernel, a C library, and a collection of userland applications, some produced by third-party free



software projects, but with a large component of in-house code. BS is arranged and configured in a distinctive fashion, optimized for use in retailing and agriculture. Downstream, Blue Sock is modified by Cloudiness Worldwide, a “virtual personal server” vendor to aquaculturists everywhere. CW may apply patches to packages within the Blue Sock distribution it offers to CW customers that conflict with maintenance directly applied by BS upstream, or which may interact with BS’s in-house code in unexpected ways. Resulting breakage in Cloudiness VPSs may be wrongly attributed by end users to quality failures at Blue Sock. If Blue Sock is a commercial distribution, this risk may well be sufficiently grave to justify trademark policy restrictions that would prevent Cloudiness Worldwide from calling its VPS operating system “Blue Sock” unless, for example, BS is accorded an opportunity to review and coordinate its technical maintenance activities with CW’s.

Communities and Companies

The difficulties that arise from overlapping rights occur most sharply at the transitions between non-commercial and commercial parties. For non-commercial projects, packagers, or distribution-makers—who as a general rule do not license rights for revenue—license terms on trademarks are set by the need to affirm quality and provenance. Commercial parties, both downstream and upstream from non-commercial producers and packagers, on the other hand, not only may wish to monetize their ancillary rights, but also to use those rights to compete more effectively against other parties in the marketplace. As a result, the interleaved structures of nested rights created by aggregating free software programs into distributions may behave counter-intuitively when parties within the structure aggressively assert their ancillary rights.

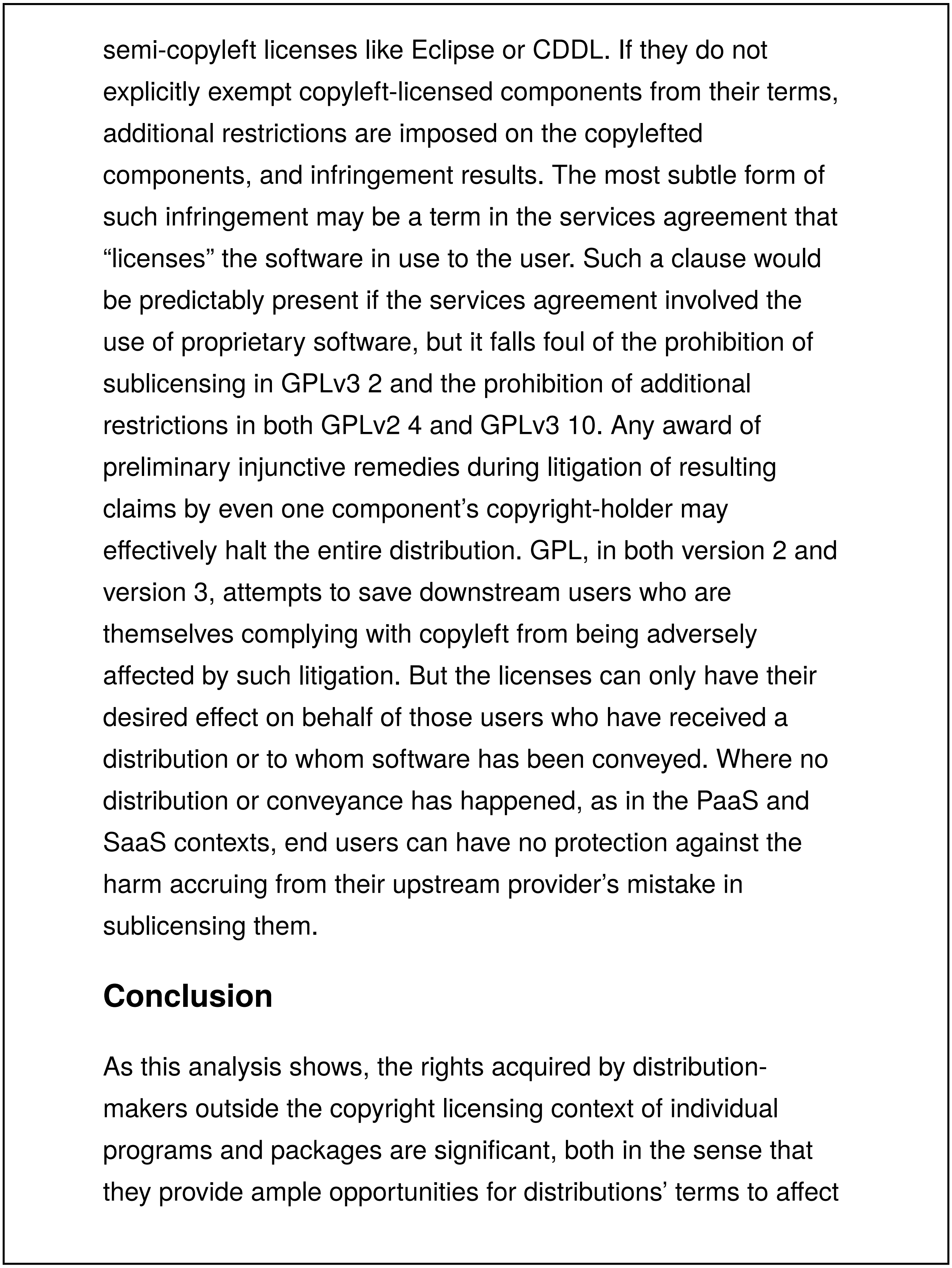
Although the same issues present at commercial/non-commercial boundaries are present everywhere, the variance in motives leads to greater chance of mutual misunderstanding and resulting social conflict.

As we have observed, copyright and trademark rules “nest” differently, in the concentric structures of free software distribution. On the trademark side, enforcement of trademark licensing policies is legally required. But non-commercial upstreams not engaged in monetizing their rights tend to have more liberal trademark licensing policies than their commercial downstreams. Thus, for example, Debian allows far more use of the Debian mark by downstream distributions than Red Hat allows. The Red Hat trademark policy does not allow even unmodified copies of RHEL to be called “Red Hat” by the distributor, without specific written permission. Debian’s policy, on the other hand, allows any factually truthful assertion of Debian’s relation to other parties’ services and software, so long as no misleading or confusing use of the Debian marks is involved. So a modified distribution of software made downstream from Canonical using modified Ubuntu packages could conceivably be called “Debian-based,” in compliance with Debian’s trademark policy, even though it could not be called “Ubuntu” under Canonical’s policy.

More generally, if non-commercial distribution A, with a liberal trademark licensing policy, is upstream from commercial distribution B, which chooses to impose more restrictions on the use of its own name in support of its business interests, the makers of downstream distributions or providers of PaaS services C, D, and E may choose to trumpet their descent from A, dropping B's name from sight altogether, to B's ultimate loss. The resulting complexities in both policy drafting and enforcement strategy have no analogues in conventional trademark practice.

The copyright rules applicable to the same situation, on the other hand, are determined by the right/left valance of the licensing of each component work. If, for example, Canonical has applied a EULA that imposes requirements on binaries made from permissively-licensed free software programs distributed as part of Ubuntu and also in Debian, those binaries cannot be copied into another distribution under non-Canonical terms. Component packagers who use permissive licenses may be surprised to discover that the terms on which downstream users receive their unmodified packages may not be the ones they themselves applied. Executables of programs subject to strong copyleft cannot be similarly restricted, because the copyleft on the binaries prohibits the imposition of any additional restrictions.

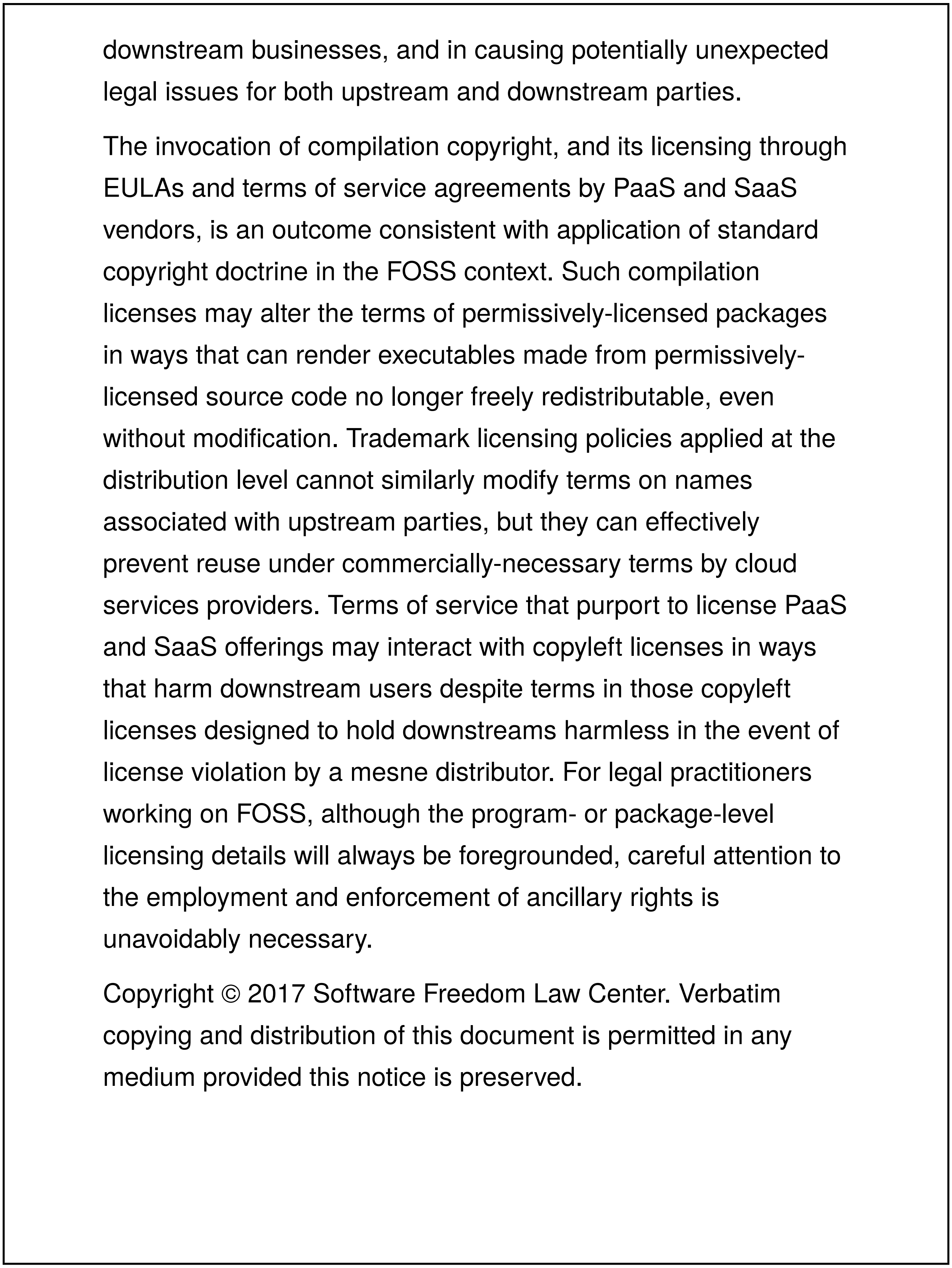
Further problems can arise from the enclosing EULAs, or services agreements, covering PaaS or SaaS distributions of FOSS stacks. The EULA or terms of service agreement may replace or supplement the terms of every permissive license used at the package or program level within the distribution, and the terms on every binary produced from source code under



semi-copyleft licenses like Eclipse or CDDL. If they do not explicitly exempt copyleft-licensed components from their terms, additional restrictions are imposed on the copylefted components, and infringement results. The most subtle form of such infringement may be a term in the services agreement that “licenses” the software in use to the user. Such a clause would be predictably present if the services agreement involved the use of proprietary software, but it falls foul of the prohibition of sublicensing in GPLv3 2 and the prohibition of additional restrictions in both GPLv2 4 and GPLv3 10. Any award of preliminary injunctive remedies during litigation of resulting claims by even one component’s copyright-holder may effectively halt the entire distribution. GPL, in both version 2 and version 3, attempts to save downstream users who are themselves complying with copyleft from being adversely affected by such litigation. But the licenses can only have their desired effect on behalf of those users who have received a distribution or to whom software has been conveyed. Where no distribution or conveyance has happened, as in the PaaS and SaaS contexts, end users can have no protection against the harm accruing from their upstream provider’s mistake in sublicensing them.

Conclusion

As this analysis shows, the rights acquired by distribution-makers outside the copyright licensing context of individual programs and packages are significant, both in the sense that they provide ample opportunities for distributions’ terms to affect

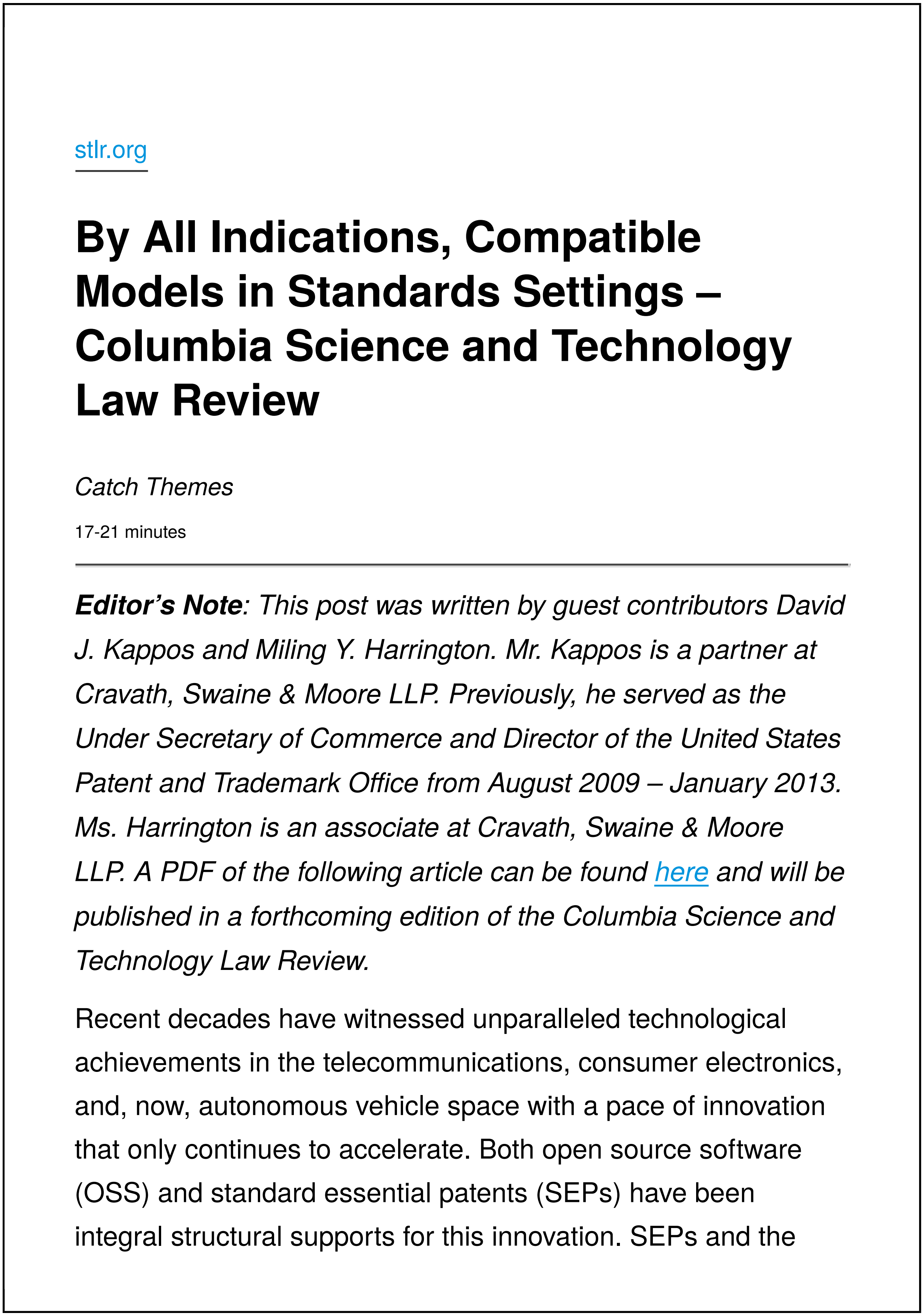


downstream businesses, and in causing potentially unexpected legal issues for both upstream and downstream parties.

The invocation of compilation copyright, and its licensing through EULAs and terms of service agreements by PaaS and SaaS vendors, is an outcome consistent with application of standard copyright doctrine in the FOSS context. Such compilation licenses may alter the terms of permissively-licensed packages in ways that can render executables made from permissively-licensed source code no longer freely redistributable, even without modification. Trademark licensing policies applied at the distribution level cannot similarly modify terms on names associated with upstream parties, but they can effectively prevent reuse under commercially-necessary terms by cloud services providers. Terms of service that purport to license PaaS and SaaS offerings may interact with copyleft licenses in ways that harm downstream users despite terms in those copyleft licenses designed to hold downstreams harmless in the event of license violation by a mesne distributor. For legal practitioners working on FOSS, although the program- or package-level licensing details will always be foregrounded, careful attention to the employment and enforcement of ancillary rights is unavoidably necessary.

Chapter 26

The Truth About OSS-FRAND (Kappos and Harrington)



stlr.org

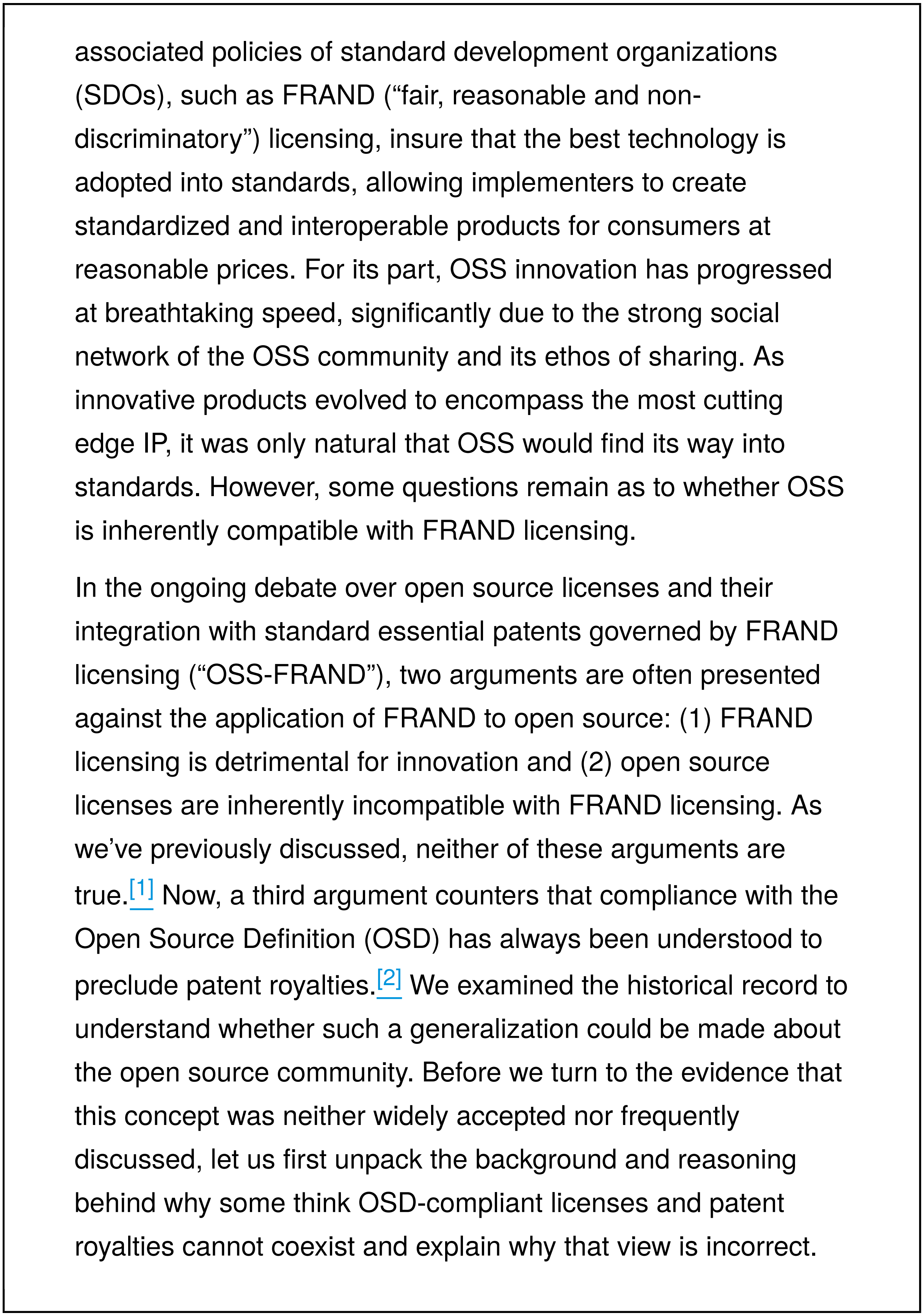
By All Indications, Compatible Models in Standards Settings – Columbia Science and Technology Law Review

Catch Themes

17-21 minutes

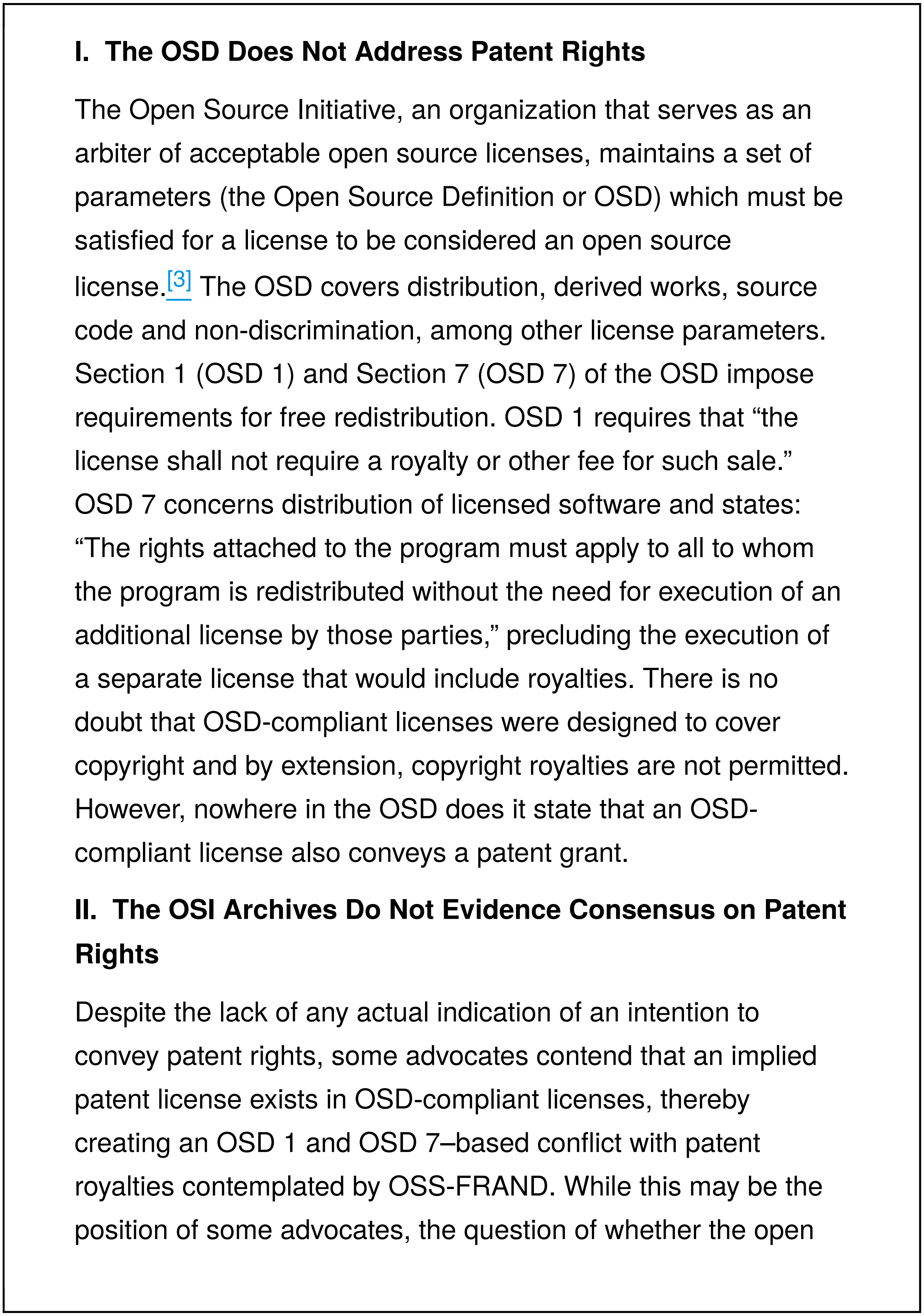
***Editor's Note**: This post was written by guest contributors David J. Kappos and Miling Y. Harrington. Mr. Kappos is a partner at Cravath, Swaine & Moore LLP. Previously, he served as the Under Secretary of Commerce and Director of the United States Patent and Trademark Office from August 2009 – January 2013. Ms. Harrington is an associate at Cravath, Swaine & Moore LLP. A PDF of the following article can be found here and will be published in a forthcoming edition of the Columbia Science and Technology Law Review.*

Recent decades have witnessed unparalleled technological achievements in the telecommunications, consumer electronics, and, now, autonomous vehicle space with a pace of innovation that only continues to accelerate. Both open source software (OSS) and standard essential patents (SEPs) have been integral structural supports for this innovation. SEPs and the



associated policies of standard development organizations (SDOs), such as FRAND (“fair, reasonable and non-discriminatory”) licensing, insure that the best technology is adopted into standards, allowing implementers to create standardized and interoperable products for consumers at reasonable prices. For its part, OSS innovation has progressed at breathtaking speed, significantly due to the strong social network of the OSS community and its ethos of sharing. As innovative products evolved to encompass the most cutting edge IP, it was only natural that OSS would find its way into standards. However, some questions remain as to whether OSS is inherently compatible with FRAND licensing.

In the ongoing debate over open source licenses and their integration with standard essential patents governed by FRAND licensing (“OSS-FRAND”), two arguments are often presented against the application of FRAND to open source: (1) FRAND licensing is detrimental for innovation and (2) open source licenses are inherently incompatible with FRAND licensing. As we’ve previously discussed, neither of these arguments are true.[1] Now, a third argument counters that compliance with the Open Source Definition (OSD) has always been understood to preclude patent royalties.[2] We examined the historical record to understand whether such a generalization could be made about the open source community. Before we turn to the evidence that this concept was neither widely accepted nor frequently discussed, let us first unpack the background and reasoning behind why some think OSD-compliant licenses and patent royalties cannot coexist and explain why that view is incorrect.

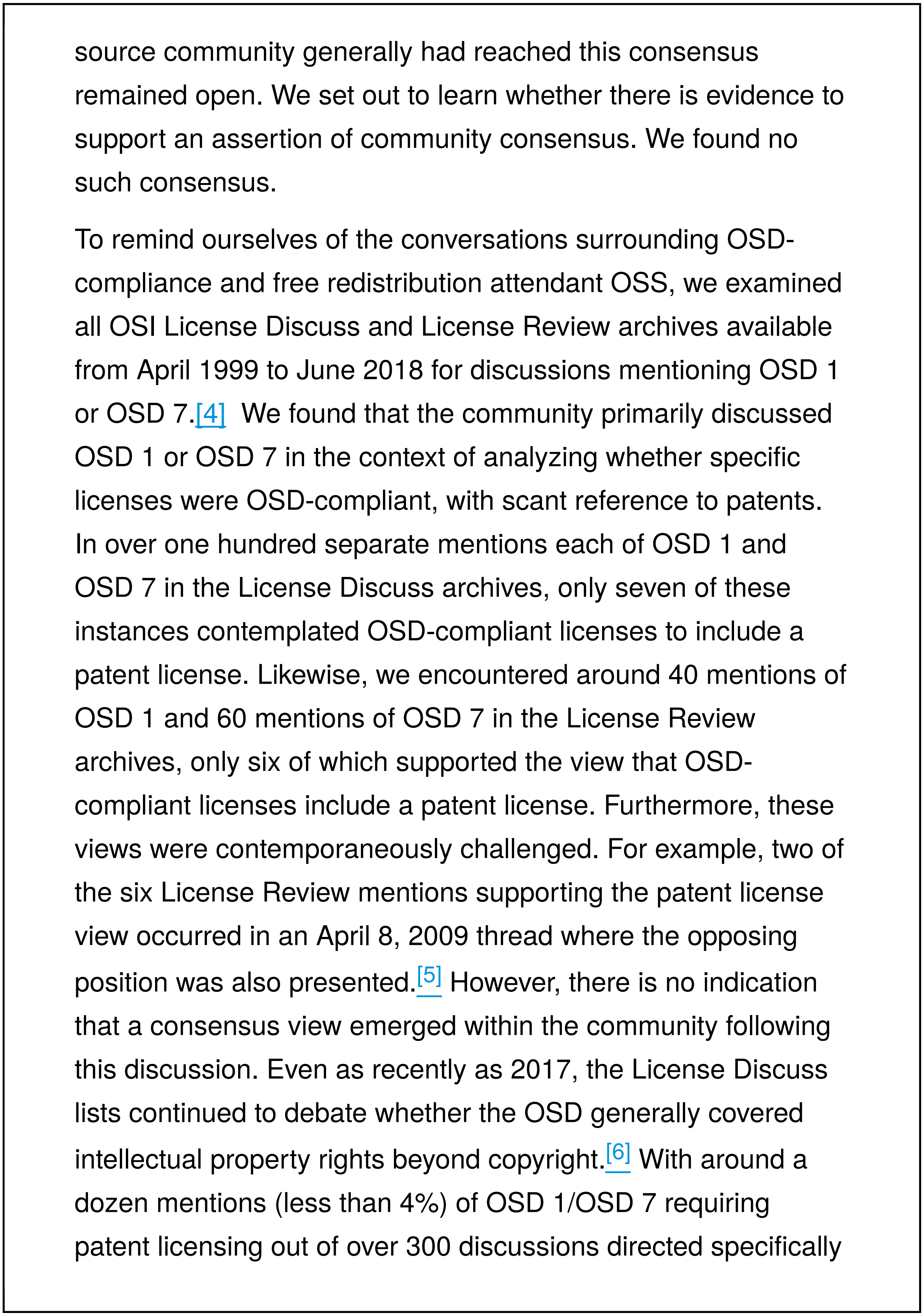


I. The OSD Does Not Address Patent Rights

The Open Source Initiative, an organization that serves as an arbiter of acceptable open source licenses, maintains a set of parameters (the Open Source Definition or OSD) which must be satisfied for a license to be considered an open source license.[3] The OSD covers distribution, derived works, source code and non-discrimination, among other license parameters. Section 1 (OSD 1) and Section 7 (OSD 7) of the OSD impose requirements for free redistribution. OSD 1 requires that "the license shall not require a royalty or other fee for such sale." OSD 7 concerns distribution of licensed software and states: "The rights attached to the program must apply to all to whom the program is redistributed without the need for execution of an additional license by those parties," precluding the execution of a separate license that would include royalties. There is no doubt that OSD-compliant licenses were designed to cover copyright and by extension, copyright royalties are not permitted. However, nowhere in the OSD does it state that an OSD-compliant license also conveys a patent grant.

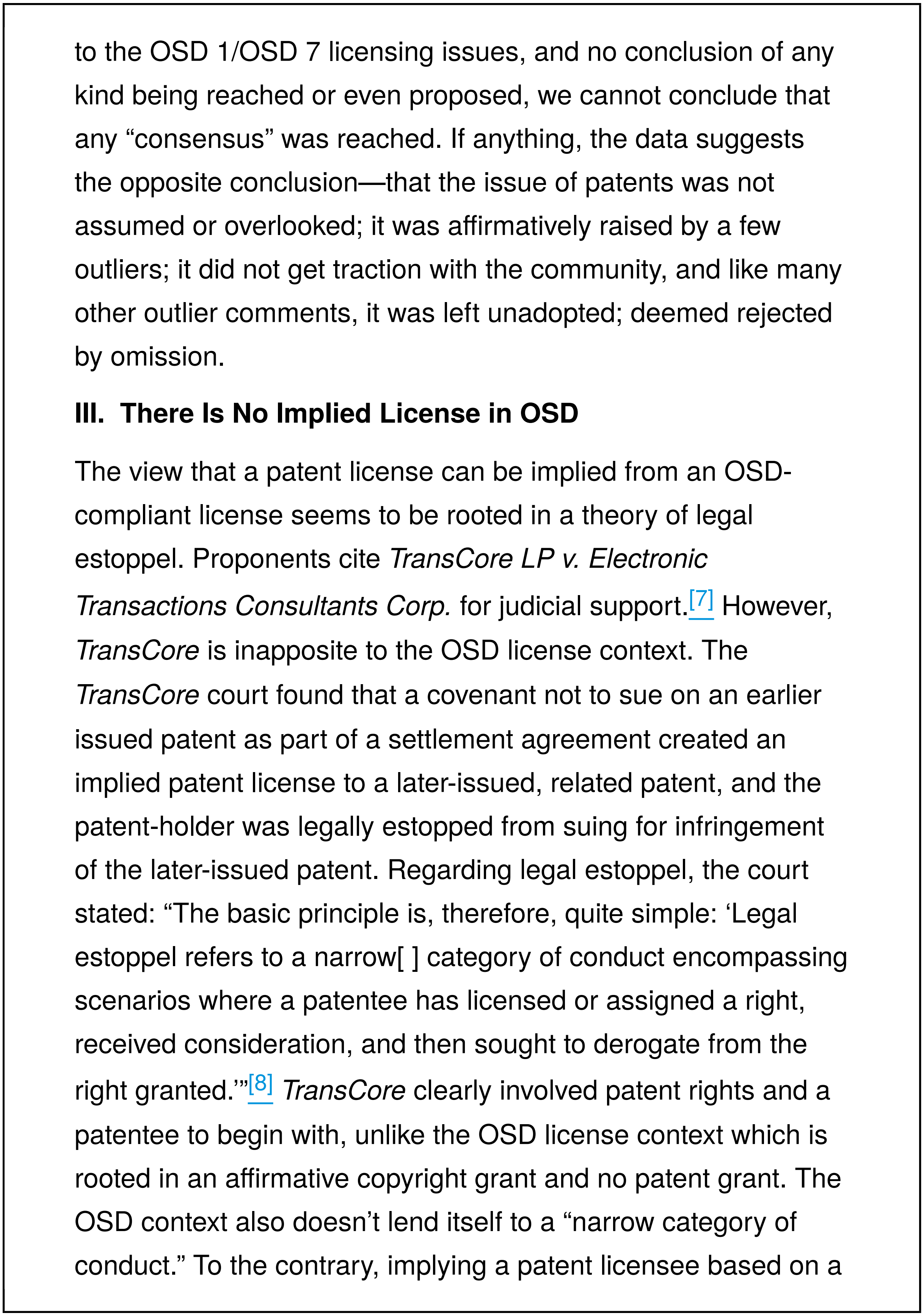
II. The OSI Archives Do Not Evidence Consensus on Patent Rights

Despite the lack of any actual indication of an intention to convey patent rights, some advocates contend that an implied patent license exists in OSD-compliant licenses, thereby creating an OSD 1 and OSD 7–based conflict with patent royalties contemplated by OSS-FRAND. While this may be the position of some advocates, the question of whether the open



source community generally had reached this consensus remained open. We set out to learn whether there is evidence to support an assertion of community consensus. We found no such consensus.

To remind ourselves of the conversations surrounding OSD-compliance and free redistribution attendant OSS, we examined all OSI License Discuss and License Review archives available from April 1999 to June 2018 for discussions mentioning OSD 1 or OSD 7.[4] We found that the community primarily discussed OSD 1 or OSD 7 in the context of analyzing whether specific licenses were OSD-compliant, with scant reference to patents. In over one hundred separate mentions each of OSD 1 and OSD 7 in the License Discuss archives, only seven of these instances contemplated OSD-compliant licenses to include a patent license. Likewise, we encountered around 40 mentions of OSD 1 and 60 mentions of OSD 7 in the License Review archives, only six of which supported the view that OSD-compliant licenses include a patent license. Furthermore, these views were contemporaneously challenged. For example, two of the six License Review mentions supporting the patent license view occurred in an April 8, 2009 thread where the opposing position was also presented.[5] However, there is no indication that a consensus view emerged within the community following this discussion. Even as recently as 2017, the License Discuss lists continued to debate whether the OSD generally covered intellectual property rights beyond copyright.[6] With around a dozen mentions (less than 4%) of OSD 1/OSD 7 requiring patent licensing out of over 300 discussions directed specifically

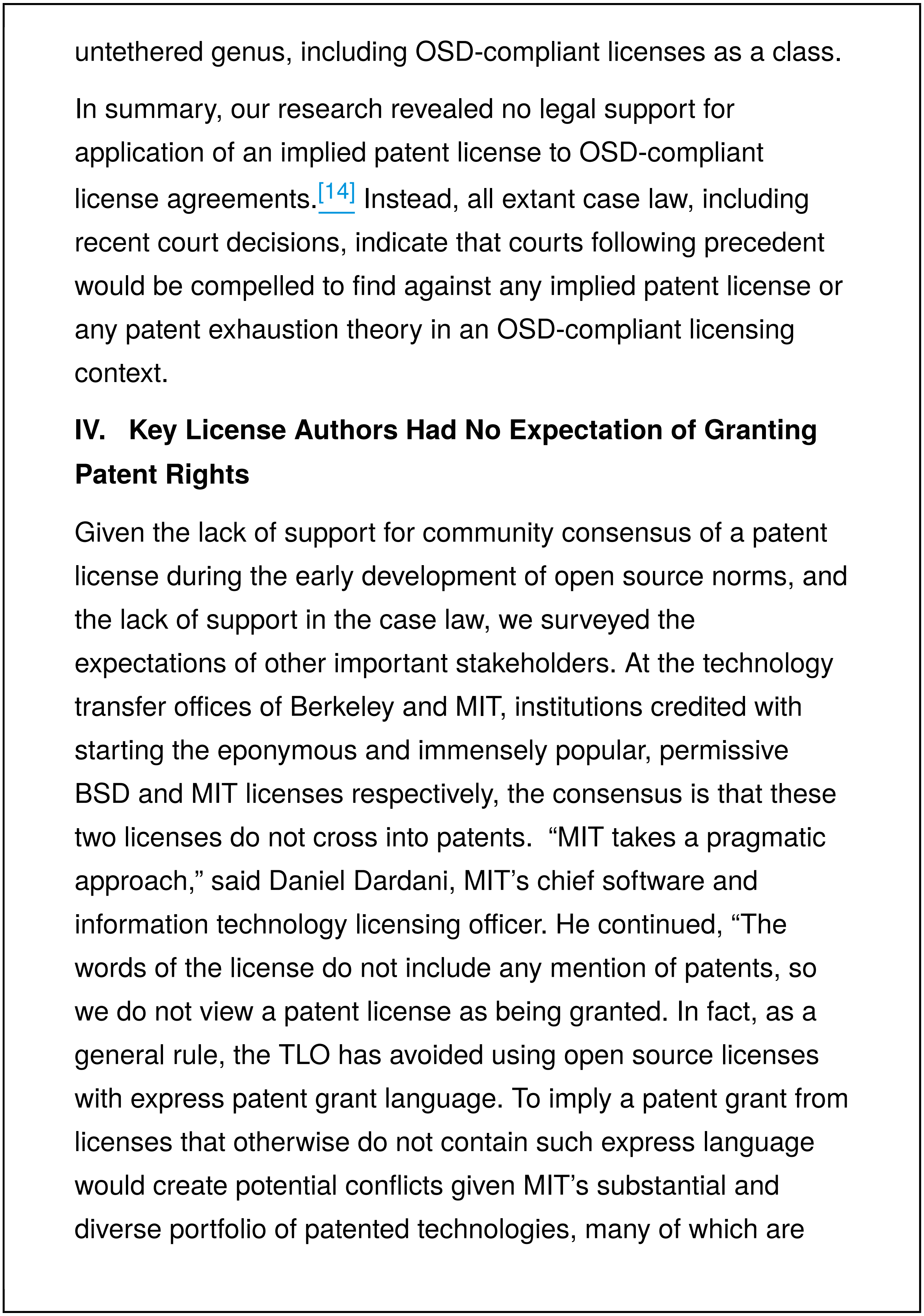


to the OSD 1/OSD 7 licensing issues, and no conclusion of any kind being reached or even proposed, we cannot conclude that any "consensus" was reached. If anything, the data suggests the opposite conclusion—that the issue of patents was not assumed or overlooked; it was affirmatively raised by a few outliers; it did not get traction with the community, and like many other outlier comments, it was left unadopted; deemed rejected by omission.

III. There Is No Implied License in OSD

The view that a patent license can be implied from an OSD-compliant license seems to be rooted in a theory of legal estoppel. Proponents cite *TransCore LP v. Electronic Transactions Consultants Corp.* for judicial support.[7] However, *TransCore* is inapposite to the OSD license context. The *TransCore* court found that a covenant not to sue on an earlier issued patent as part of a settlement agreement created an implied patent license to a later-issued, related patent, and the patent-holder was legally estopped from suing for infringement of the later-issued patent. Regarding legal estoppel, the court stated: "The basic principle is, therefore, quite simple: 'Legal estoppel refers to a narrow[ ] category of conduct encompassing scenarios where a patentee has licensed or assigned a right, received consideration, and then sought to derogate from the right granted.'"[8] *TransCore* clearly involved patent rights and a patentee to begin with, unlike the OSD license context which is rooted in an affirmative copyright grant and no patent grant. The OSD context also doesn't lend itself to a "narrow category of conduct." To the contrary, implying a patent licensee based on a free, unsigned, automatic copyright license would sweep in a broad array of conduct. Furthermore, although the Federal Circuit discussed legal estoppel in *Wang Laboratories, Inc. v. Mitsubishi Electronics. America, Inc.*, 103 F.3d 1571, (Fed. Cir. 1997), its ultimate finding of an implied patent license was rooted in equitable rather than legal estoppel.[9] While legal estoppel analysis looks for “an affirmative grant of consent or permission to make, use, or sell; i.e. a license,” equitable estoppel analysis “focuses on ‘misleading’ conduct suggesting that the patentee will not enforce patent rights.”[10] Equitable estoppel has even less of a basis to be applied broadly to OSD licenses as a class.

Our research was unsuccessful in finding any court case that has considered whether patent licenses are implied by open source licenses in the absence of express language. But the caselaw surrounding implied licenses indicates that courts are hesitant to imply a license where one is not expressly set forth. The District of Delaware quoted the Federal Circuit in the recent case *Endo Pharmaceuticals Inc. v. Amneal Pharmaceuticals, LLC*, 224 F.Supp.3d 368 (D. Del. 2016), “[J]udicially implied licenses are rare under any doctrine,” in concluding that defendant Teva had not demonstrated facts supporting an implied patent license.[11] Likewise, the Northern District of California has stated, “’Courts have found implied licenses only in narrow circumstances where one party created a work at [the other’s] request and handed it over, intending that [the other] copy and distribute it.’”[12] The implied patent license inquiry in general is narrow and fact-specific,[13] and thus unsuited to any



untethered genus, including OSD-compliant licenses as a class.

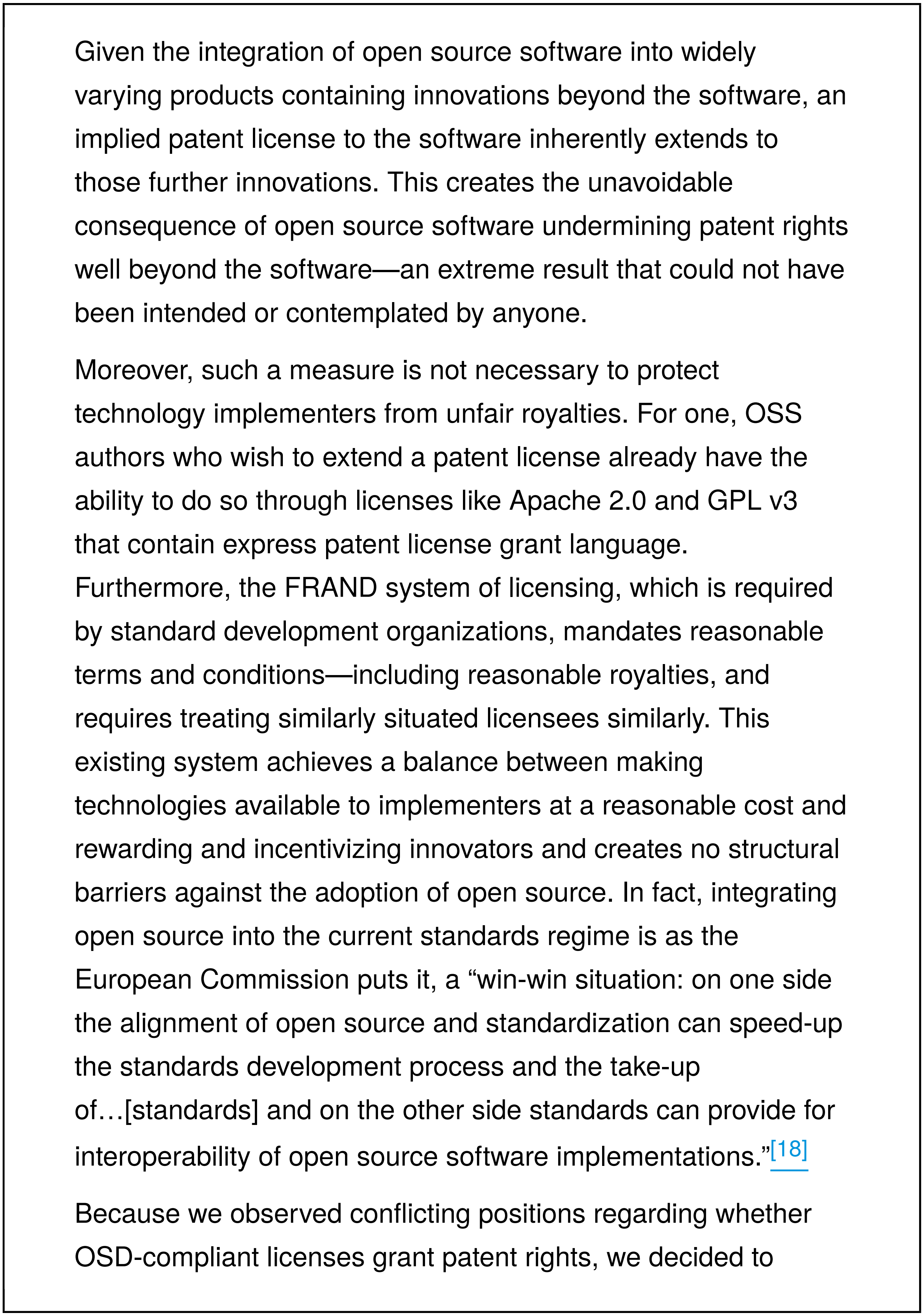
In summary, our research revealed no legal support for application of an implied patent license to OSD-compliant license agreements.[14] Instead, all extant case law, including recent court decisions, indicate that courts following precedent would be compelled to find against any implied patent license or any patent exhaustion theory in an OSD-compliant licensing context.

IV. Key License Authors Had No Expectation of Granting Patent Rights

Given the lack of support for community consensus of a patent license during the early development of open source norms, and the lack of support in the case law, we surveyed the expectations of other important stakeholders. At the technology transfer offices of Berkeley and MIT, institutions credited with starting the eponymous and immensely popular, permissive BSD and MIT licenses respectively, the consensus is that these two licenses do not cross into patents. “MIT takes a pragmatic approach,” said Daniel Dardani, MIT’s chief software and information technology licensing officer. He continued, “The words of the license do not include any mention of patents, so we do not view a patent license as being granted. In fact, as a general rule, the TLO has avoided using open source licenses with express patent grant language. To imply a patent grant from licenses that otherwise do not contain such express language would create potential conflicts given MIT’s substantial and diverse portfolio of patented technologies, many of which are exclusively licensed to companies."[15] This position is shared by Berkeley's Office of Technology Licensing (OTL). Curt Theisen, the Associate Director of the OTL, adds, "The Berkeley OTL has never taken the position that the BSD includes a patent grant. In fact, we regularly advise our community members that the BSD license is an excellent OSS license to use because it permits broad licensing of software with minimal restrictions and maximum compatibility with other software and licenses."[16] Both Berkeley's and MIT's views fit into the broader consensus that permissive licenses, unlike copyleft licenses, do not contain restrictive language and are compatible with FRAND licensing.[17] We are thus compelled to conclude that the view of certain OSD-compliant licenses necessarily granting patent rights causing incompatibility with FRAND is neither rooted in the past nor serves the interests of the present.

V. A Forced OSS-FRAND Free Patent License Disturbs the Innovation Ecosystem

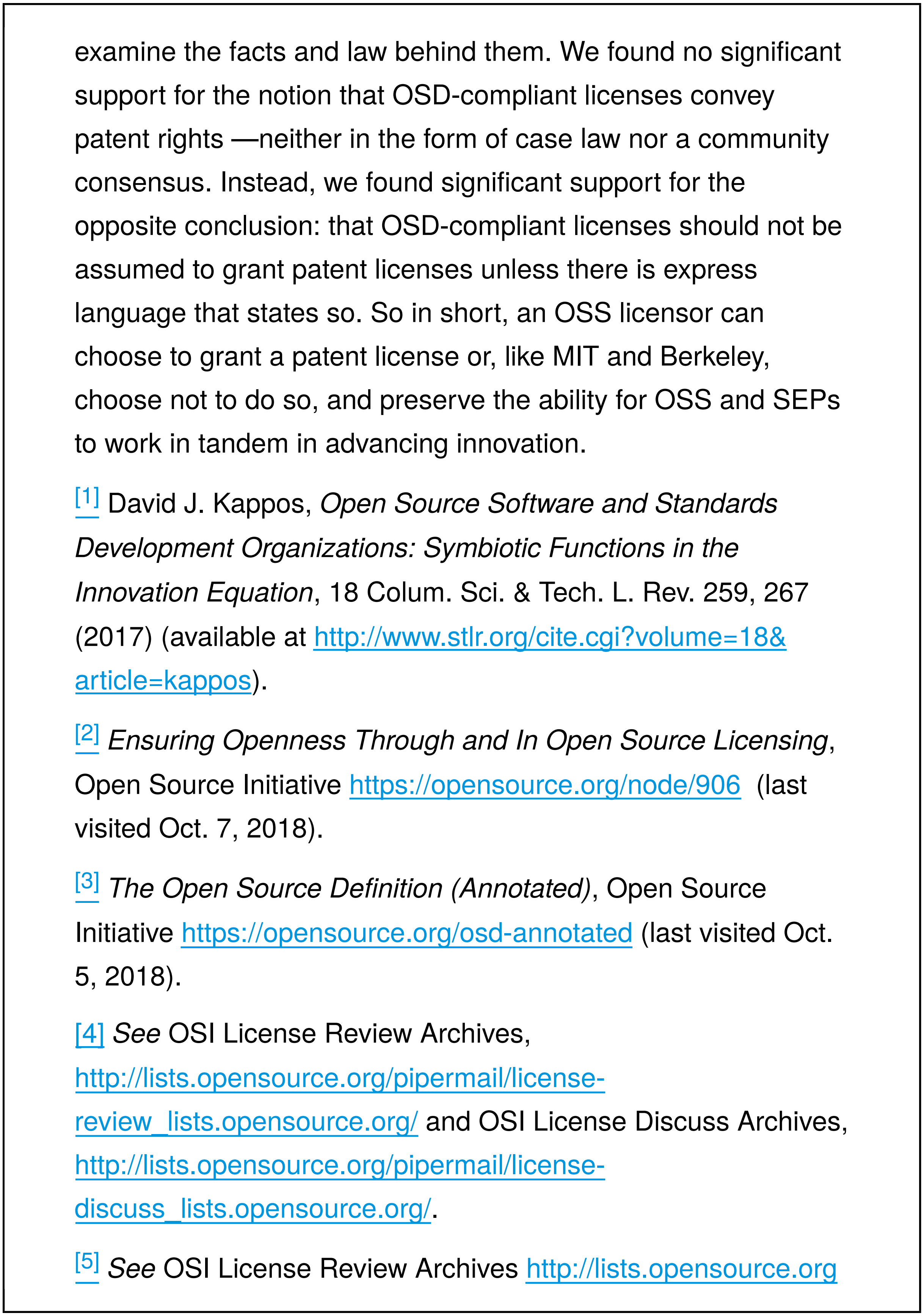
Turning finally to the bigger picture, it is important to understand that OSD-compliant licenses in the context of OSS-FRAND cannot be examined in isolation. The software they cover is integrated into highly sophisticated products (such as smart phones) that encompass intellectual property covering myriad functions and components. To declare OSD-compliant licenses to be incompatible with patent royalties both over-extends the reach of the software license to functions and components beyond the scope of the license, and "solves" a problem that is already amply addressed by existing safeguards.



Given the integration of open source software into widely varying products containing innovations beyond the software, an implied patent license to the software inherently extends to those further innovations. This creates the unavoidable consequence of open source software undermining patent rights well beyond the software—an extreme result that could not have been intended or contemplated by anyone.

Moreover, such a measure is not necessary to protect technology implementers from unfair royalties. For one, OSS authors who wish to extend a patent license already have the ability to do so through licenses like Apache 2.0 and GPL v3 that contain express patent license grant language. Furthermore, the FRAND system of licensing, which is required by standard development organizations, mandates reasonable terms and conditions—including reasonable royalties, and requires treating similarly situated licensees similarly. This existing system achieves a balance between making technologies available to implementers at a reasonable cost and rewarding and incentivizing innovators and creates no structural barriers against the adoption of open source. In fact, integrating open source into the current standards regime is as the European Commission puts it, a “win-win situation: on one side the alignment of open source and standardization can speed-up the standards development process and the take-up of…[standards] and on the other side standards can provide for interoperability of open source software implementations.”[18]

Because we observed conflicting positions regarding whether OSD-compliant licenses grant patent rights, we decided to



examine the facts and law behind them. We found no significant support for the notion that OSD-compliant licenses convey patent rights —neither in the form of case law nor a community consensus. Instead, we found significant support for the opposite conclusion: that OSD-compliant licenses should not be assumed to grant patent licenses unless there is express language that states so. So in short, an OSS licensor can choose to grant a patent license or, like MIT and Berkeley, choose not to do so, and preserve the ability for OSS and SEPs to work in tandem in advancing innovation.

[1] David J. Kappos, *Open Source Software and Standards Development Organizations: Symbiotic Functions in the Innovation Equation*, 18 Colum. Sci. & Tech. L. Rev. 259, 267 (2017) (available at http://www.stlr.org/cite.cgi?volume=18&article=kappos).

[2] *Ensuring Openness Through and In Open Source Licensing*, Open Source Initiative https://opensource.org/node/906 (last visited Oct. 7, 2018).

[3] *The Open Source Definition (Annotated)*, Open Source Initiative https://opensource.org/osd-annotated (last visited Oct. 5, 2018).

[4] *See* OSI License Review Archives, http://lists.opensource.org/pipermail/license-review_lists.opensource.org/ and OSI License Discuss Archives, http://lists.opensource.org/pipermail/license-discuss_lists.opensource.org/.

[5] *See* OSI License Review Archives http://lists.opensource.org

/pipermail/license-review_lists.opensource.org/2009-April/thread.html at April 8, 2009 ("*For approval: MXM Public license*" thread).

[6] *See* OSI License Discuss Archives, http://lists.opensource.org/pipermail/license-discuss_lists.opensource.org/2017-March/thread.html at March 7, 2017 ("*Patent rights and the OSD"* thread).

[7] Christian H. Nadan, *Closing the Loophole: Open Source Licensing & the Implied Patent License*, 26 Computer & Internet Law. 1 at 3 (Aug. 2009) (citing 563 F.3d 1271 (Fed. Cir. 2009)).

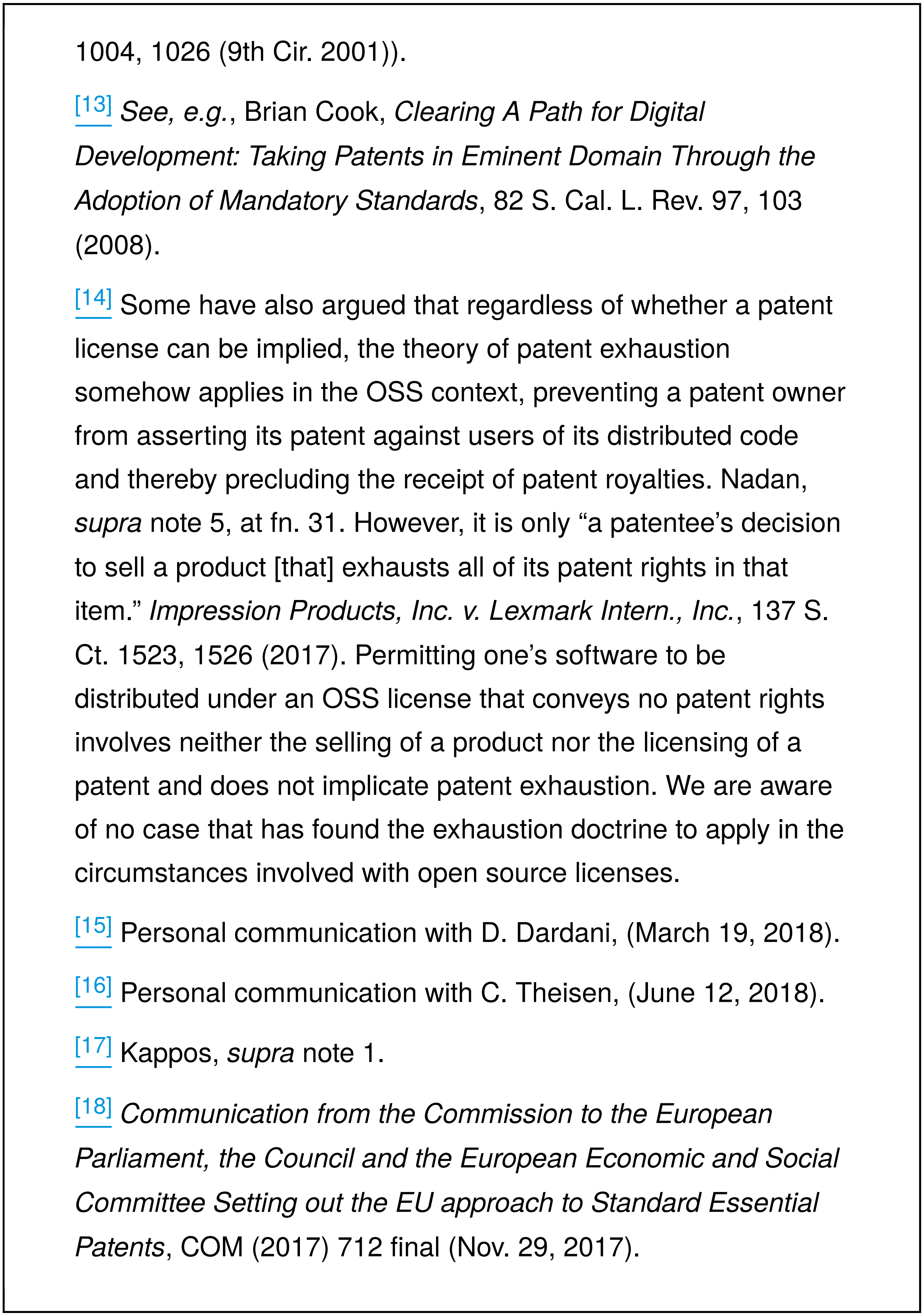
[8] 563 F. 3d at 1279 (quoting Wang Labs., Inc. v. Mitsubishi Elecs. Am., Inc., 103 F.3d 1571, 1582 (Fed. Cir. 1997))

[9] 103 F.3d at 1582 (Fed. Cir. 1997) (finding that Wang's behavior over six years including repeatedly attempting to convince Mitsubishi to join the SIMMs market, providing Mitsubishi with designs, purchasing SIMMs from Mitsubishi, lobbying for Wang's design to become an industry standard and receiving payment from Mitsubishi, were enough for Mitsubishi to infer that it had consent to use Wang's patents).

[10] *Id.* at 1581.

[11]Endo Pharm. Inc. v. Amneal Pharm, LLC, 224 F.Supp.3d 368, 382 (D. Del. 2016) (quoting *Wang Labs., Inc.*, 103 F.3d at 1581).

[12] Oracle Am., Inc. v. Terix Computer Co., Inc., No. 5:13-CV-03385 PSG, 2015 WL 2090191, at *8 (N.D. Cal. May 5, 2015) (quoting A & M Records, Inc. v. Napster, Inc., 239 F.3d



1004, 1026 (9th Cir. 2001)).

[13] *See, e.g.*, Brian Cook, *Clearing A Path for Digital Development: Taking Patents in Eminent Domain Through the Adoption of Mandatory Standards*, 82 S. Cal. L. Rev. 97, 103 (2008).

[14] Some have also argued that regardless of whether a patent license can be implied, the theory of patent exhaustion somehow applies in the OSS context, preventing a patent owner from asserting its patent against users of its distributed code and thereby precluding the receipt of patent royalties. Nadan, *supra* note 5, at fn. 31. However, it is only “a patentee’s decision to sell a product [that] exhausts all of its patent rights in that item.” *Impression Products, Inc. v. Lexmark Intern., Inc.*, 137 S. Ct. 1523, 1526 (2017). Permitting one’s software to be distributed under an OSS license that conveys no patent rights involves neither the selling of a product nor the licensing of a patent and does not implicate patent exhaustion. We are aware of no case that has found the exhaustion doctrine to apply in the circumstances involved with open source licenses.

[15] Personal communication with D. Dardani, (March 19, 2018).

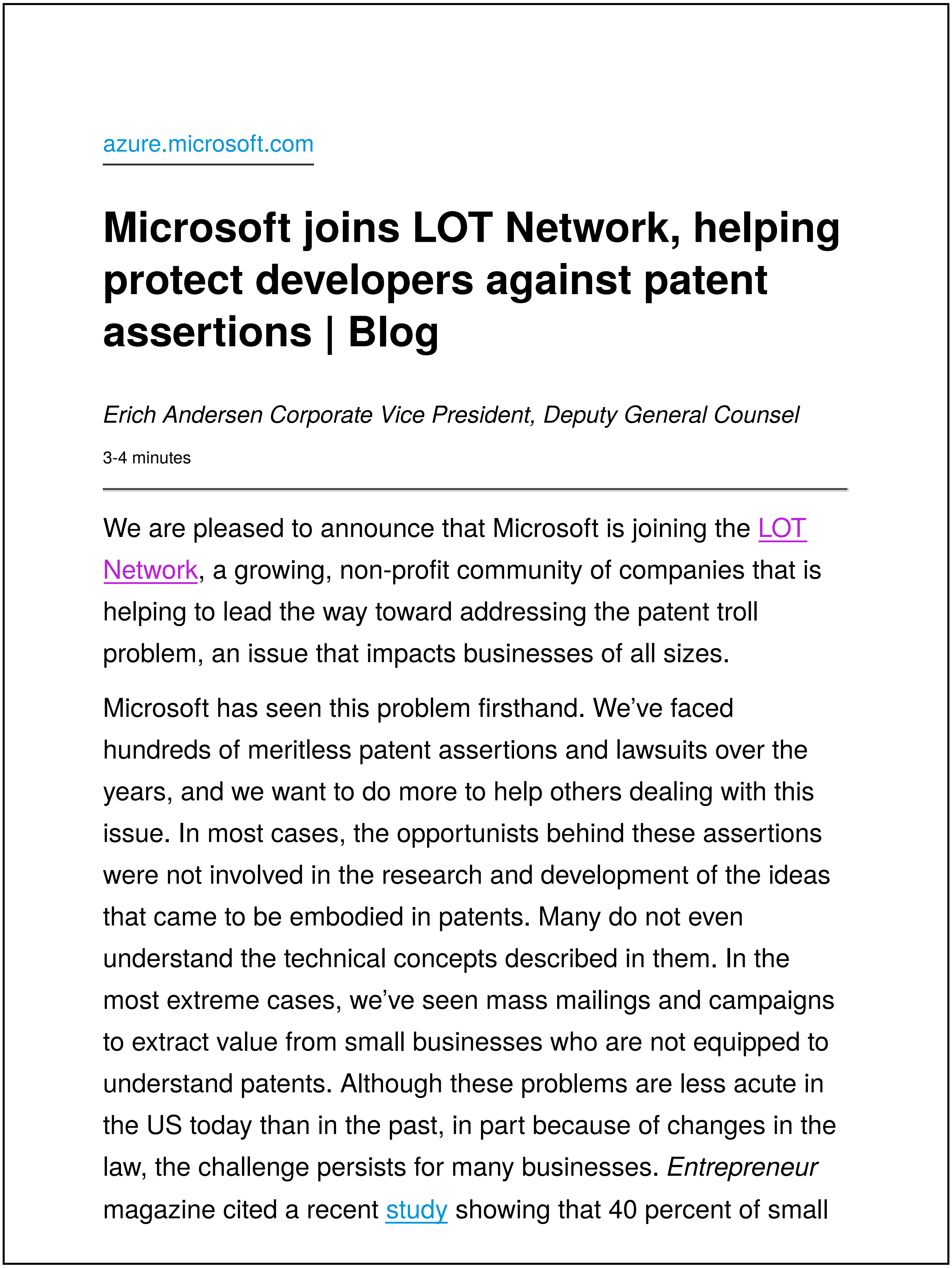
[16] Personal communication with C. Theisen, (June 12, 2018).

[17] Kappos, *supra* note 1.

[18] *Communication from the Commission to the European Parliament, the Council and the European Economic and Social Committee Setting out the EU approach to Standard Essential Patents*, COM (2017) 712 final (Nov. 29, 2017).

Chapter 27

Microsoft joins LOT Network



azure.microsoft.com

Microsoft joins LOT Network, helping protect developers against patent assertions | Blog

Erich Andersen Corporate Vice President, Deputy General Counsel

3-4 minutes

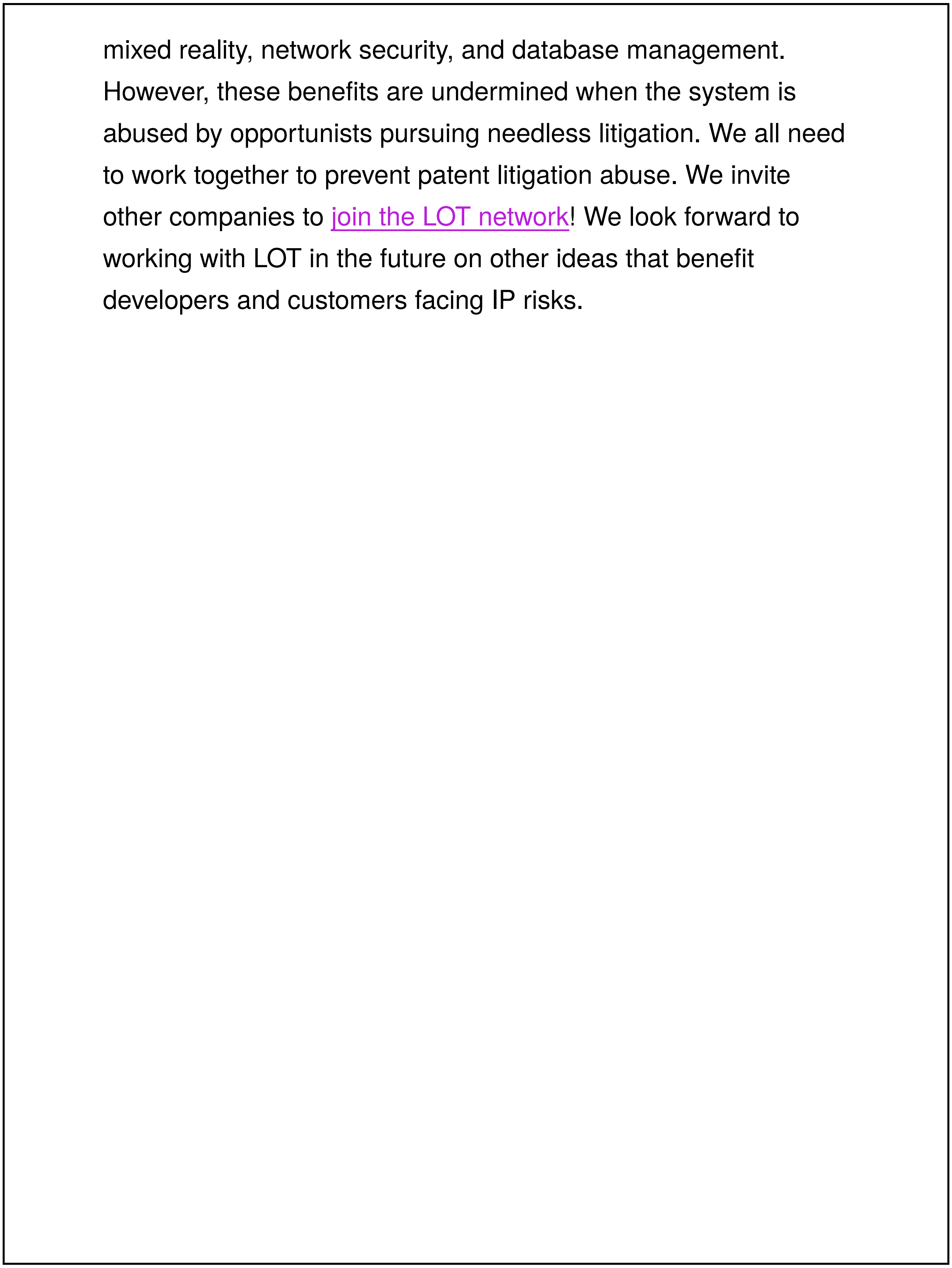
We are pleased to announce that Microsoft is joining the LOT Network, a growing, non-profit community of companies that is helping to lead the way toward addressing the patent troll problem, an issue that impacts businesses of all sizes.

Microsoft has seen this problem firsthand. We've faced hundreds of meritless patent assertions and lawsuits over the years, and we want to do more to help others dealing with this issue. In most cases, the opportunists behind these assertions were not involved in the research and development of the ideas that came to be embodied in patents. Many do not even understand the technical concepts described in them. In the most extreme cases, we've seen mass mailings and campaigns to extract value from small businesses who are not equipped to understand patents. Although these problems are less acute in the US today than in the past, in part because of changes in the law, the challenge persists for many businesses. *Entrepreneur* magazine cited a recent study showing that 40 percent of small companies involved in patent litigation reported “significant operational impact” from those suits, which some described as a “death knell.”

What does all of this mean for you if you’re a software developer or in the technology business? It means that Microsoft is taking another step to help stop patents from being asserted against you by companies running aggressive monetization campaigns. It also means that Microsoft is aligning with other industry leaders on this topic and committing to do more in the future to address IP risk. By joining the LOT network, we are committing to license our patents for free to other members if we ever transfer them to companies in the business of asserting patents. This pledge has immediate value to the nearly 300 members of the LOT community today, which covers approximately 1.35 million patents.

This also means we are continuing on the path we started with the introduction of the Azure IP Advantage program in 2017. As part of that program, Microsoft said that it would defend and indemnify developers against claims of intellectual property infringement even if the service powering Azure was built on open source. We also said that if we transferred a patent to a company in the business of asserting patents, then Azure customers would get a license for free. Our LOT membership expands this pledge to other companies in the LOT network.

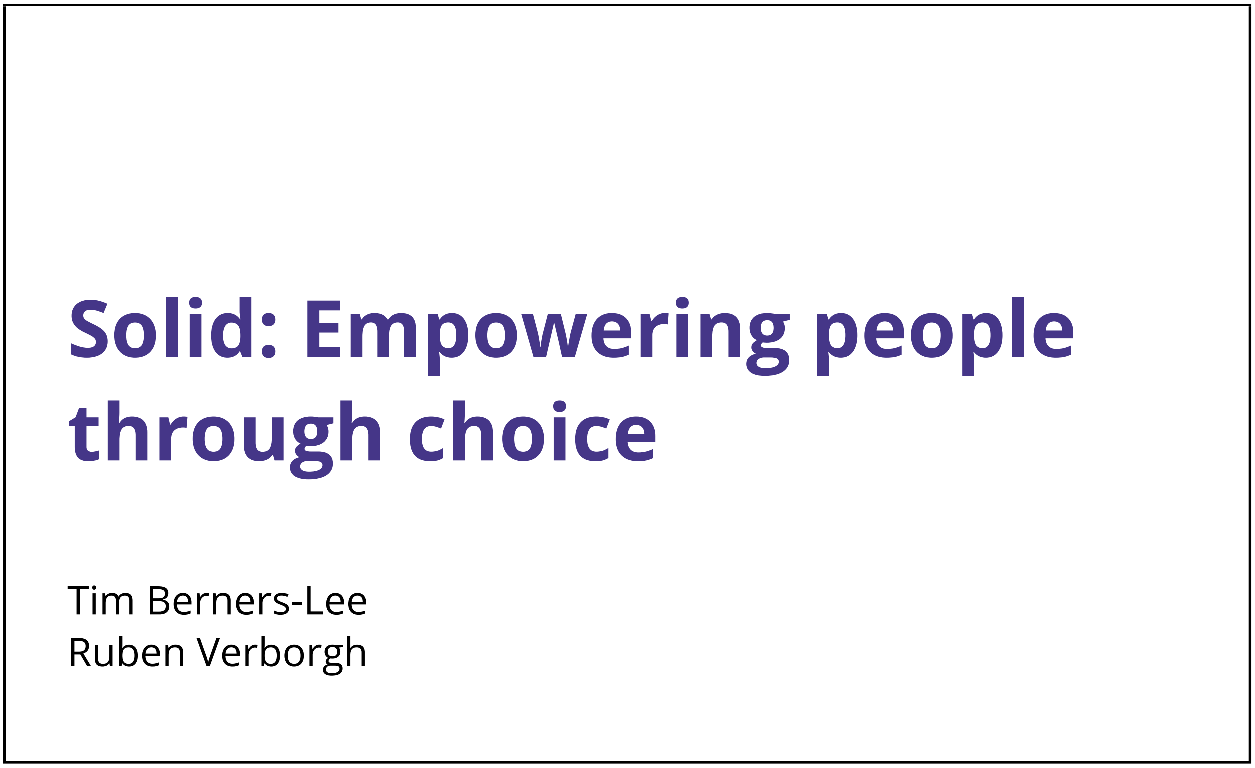
Patents and intellectual property still play an important role in our industry because they protect breakthrough innovations and allow companies large and small to recoup research and development investments in areas like artificial intelligence,



mixed reality, network security, and database management. However, these benefits are undermined when the system is abused by opportunists pursuing needless litigation. We all need to work together to prevent patent litigation abuse. We invite other companies to join the LOT network! We look forward to working with LOT in the future on other ideas that benefit developers and customers facing IP risks.

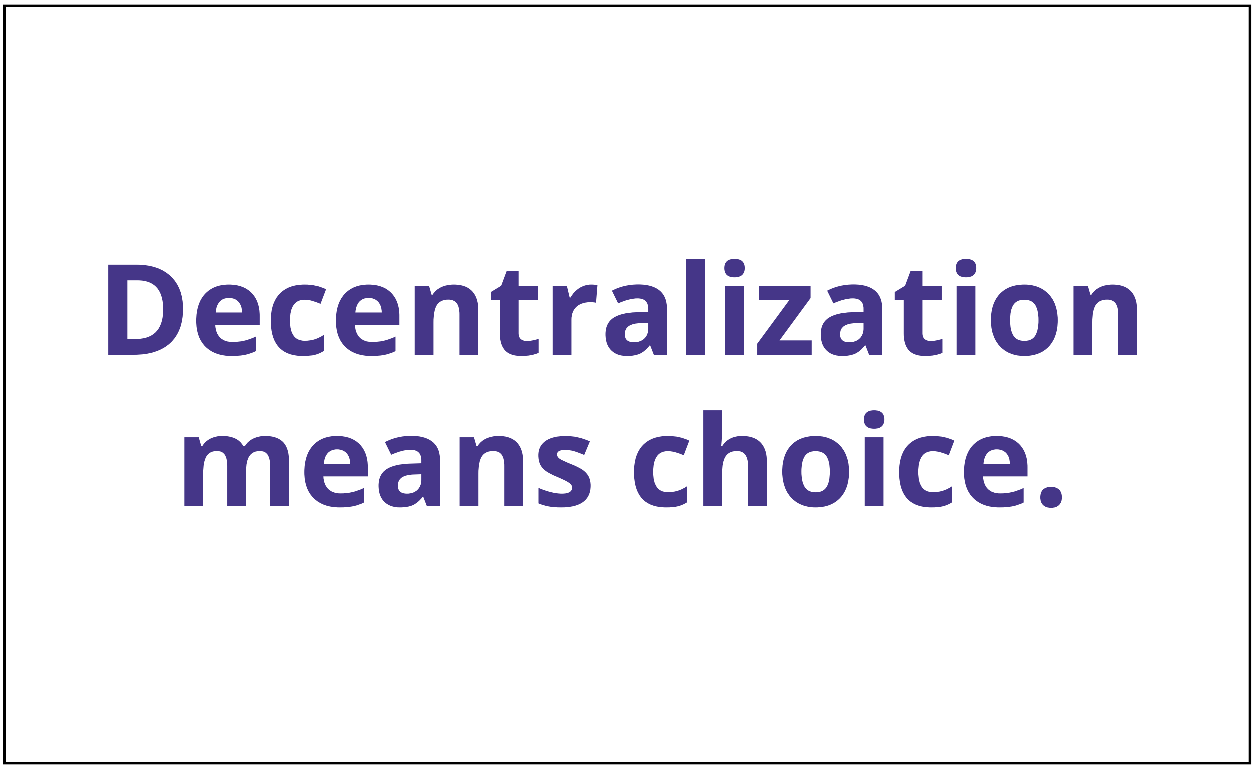
Chapter 28

Solid: Empowering People Through Choice (Berners-Lee and Verborgh)

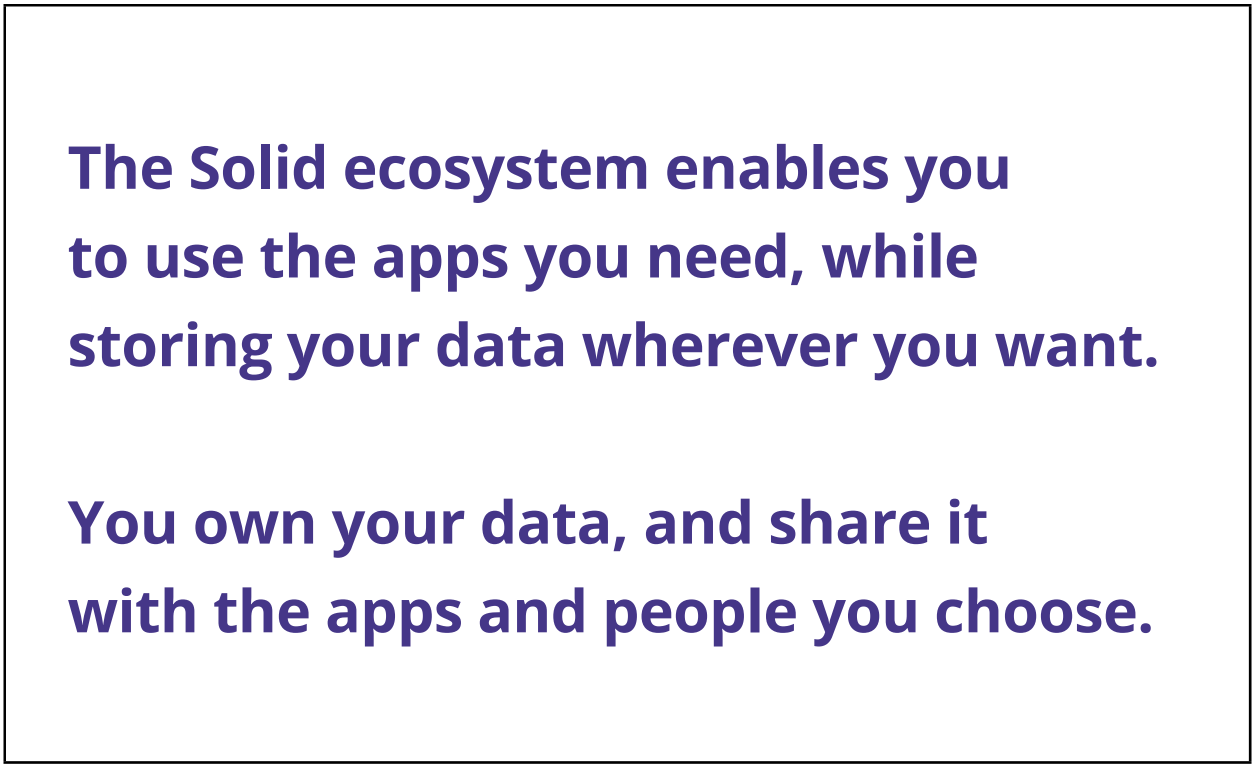


Solid: Empowering people through choice

Tim Berners-Lee
Ruben Verborgh

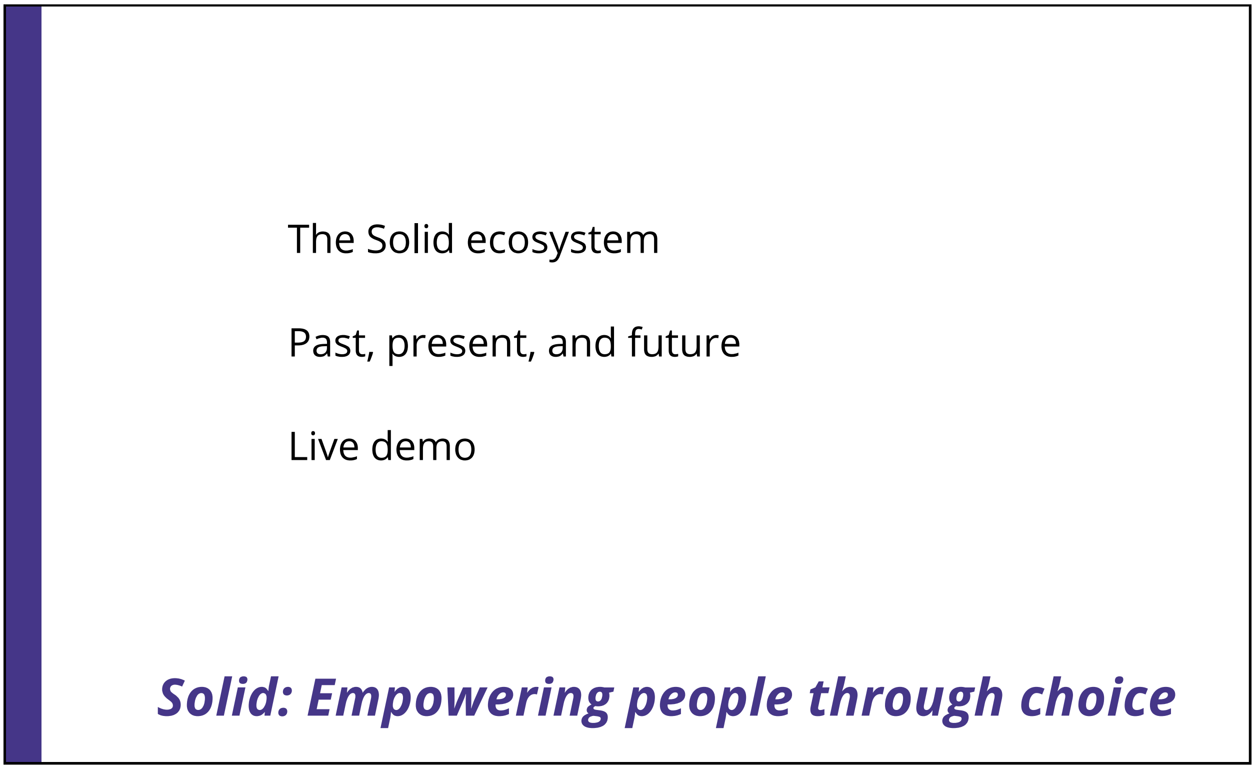


Decentralization means choice.



The Solid ecosystem enables you to use the apps you need, while storing your data wherever you want.

You own your data, and share it with the apps and people you choose.



The Solid ecosystem

Past, present, and future

Live demo

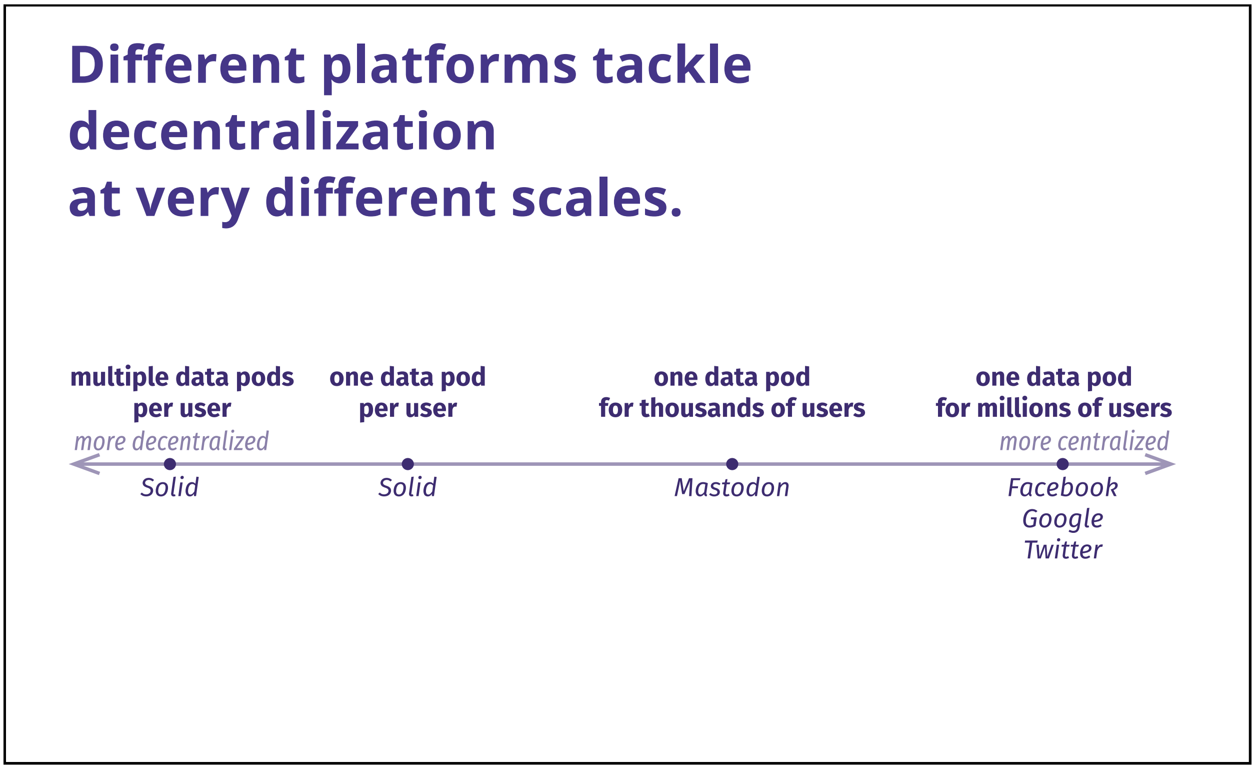
Solid: Empowering people through choice

The Solid ecosystem

Past, present, and future

Live demo

Solid: Empowering people through choice



Different platforms tackle decentralization at very different scales.

multiple data pods per user

one data pod per user

one data pod for thousands of users

one data pod for millions of users

more decentralized

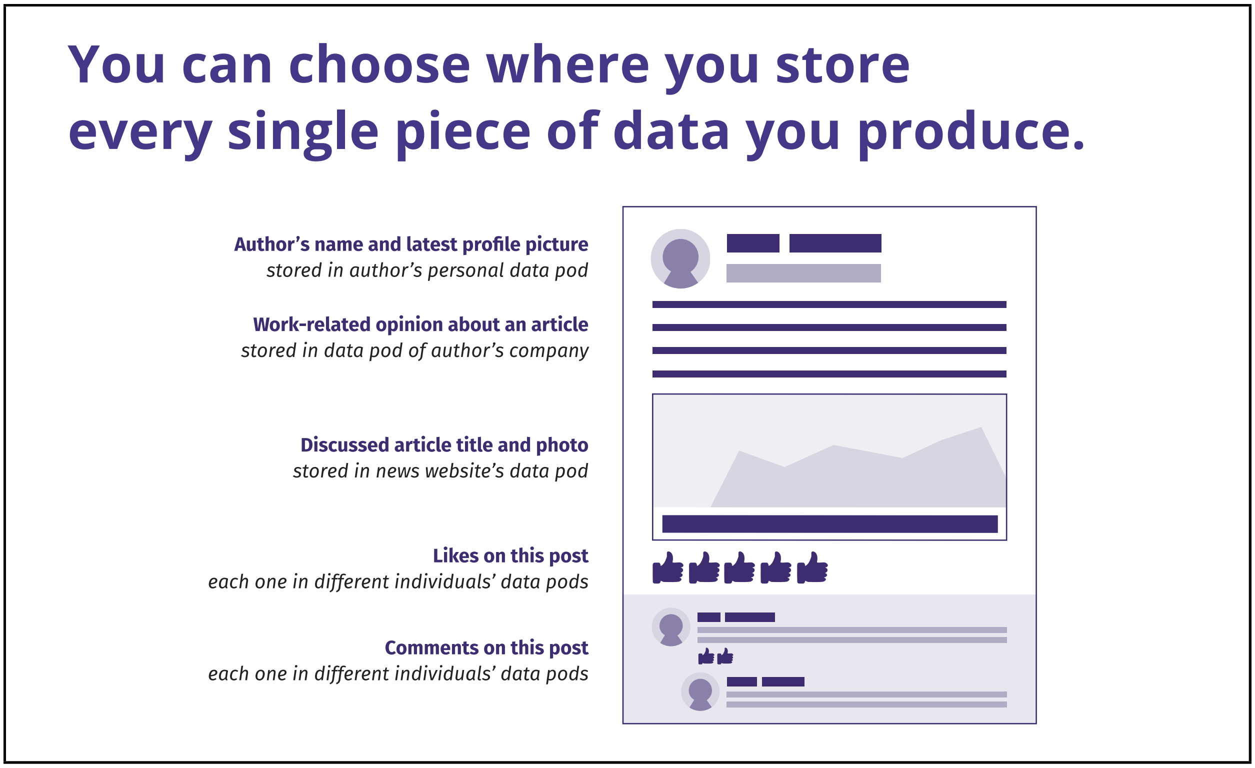
more centralized

Solid

Solid

Mastodon

Facebook
Google
Twitter



You can choose where you store every single piece of data you produce.

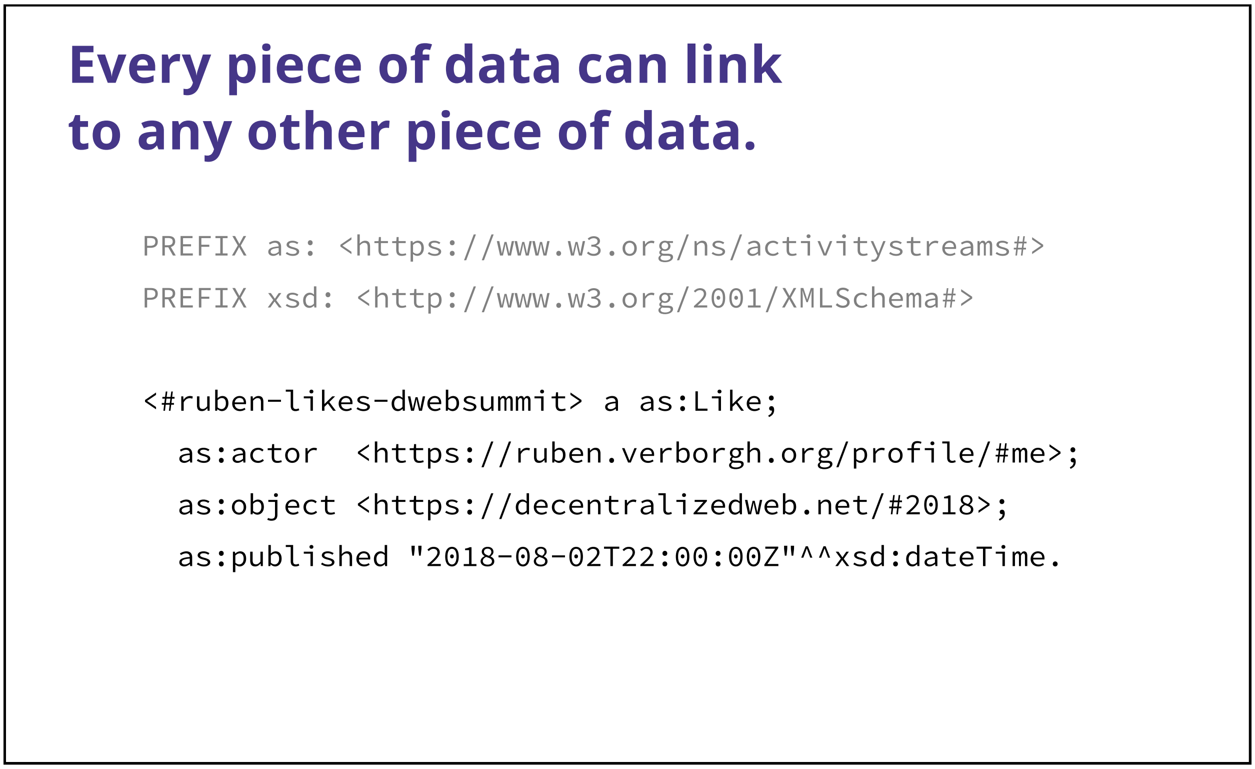
Author's name and latest profile picture
stored in author's personal data pod

Work-related opinion about an article
stored in data pod of author's company

Discussed article title and photo
stored in news website's data pod

Likes on this post
each one in different individuals' data pods

Comments on this post
each one in different individuals' data pods

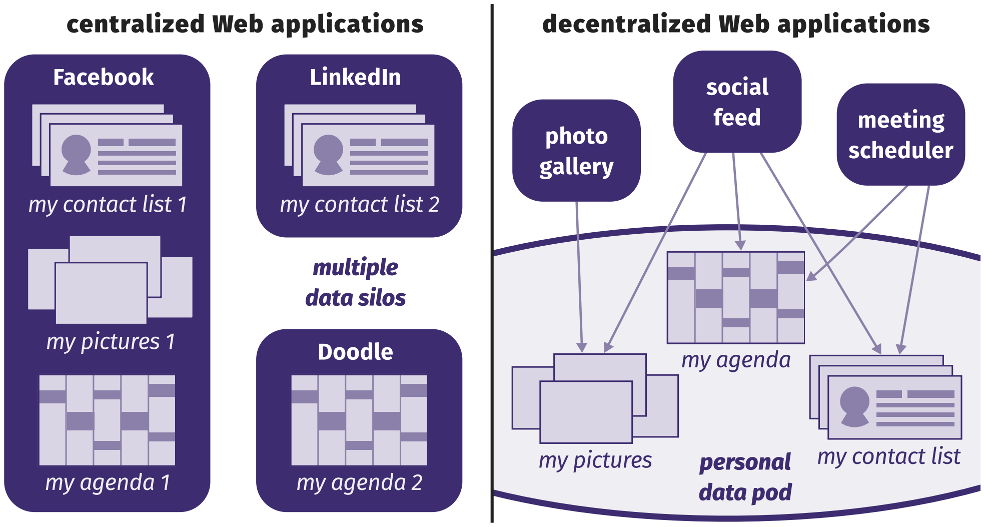


Every piece of data can link to any other piece of data.

```
PREFIX as: <https://www.w3.org/ns/activitystreams#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>

<#ruben-likes-dwebsummit> a as:Like;
  as:actor  <https://ruben.verborgh.org/profile/#me>;
  as:object <https://decentralizedweb.net/#2018>;
  as:published "2018-08-02T22:00:00Z"^^xsd:dateTime.
```

You can grant apps and people access to very specific parts of your data.



centralized Web applications

Facebook

my contact list 1

my pictures 1

my agenda 1

LinkedIn

my contact list 2

multiple data silos

Doodle

my agenda 2

decentralized Web applications

photo gallery

social feed

meeting scheduler

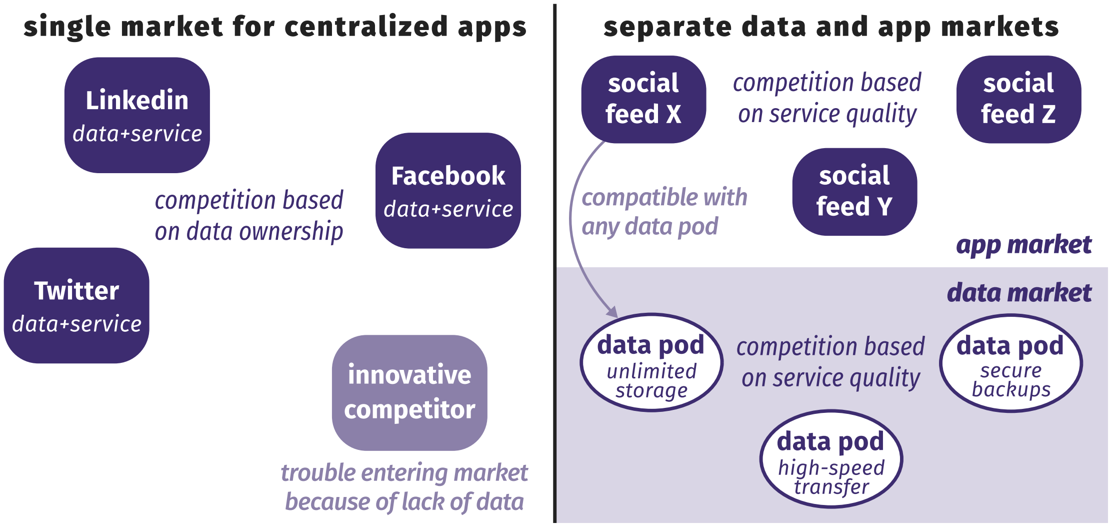
my agenda

my pictures

personal data pod

my contact list

Separating app and storage competition drives permissionless innovation.



single market for centralized apps

Linkedin
data+service

Facebook
data+service

competition based on data ownership

Twitter
data+service

innovative competitor

trouble entering market because of lack of data

separate data and app markets

social feed X

competition based on service quality

social feed Z

compatible with any data pod

social feed Y

app market

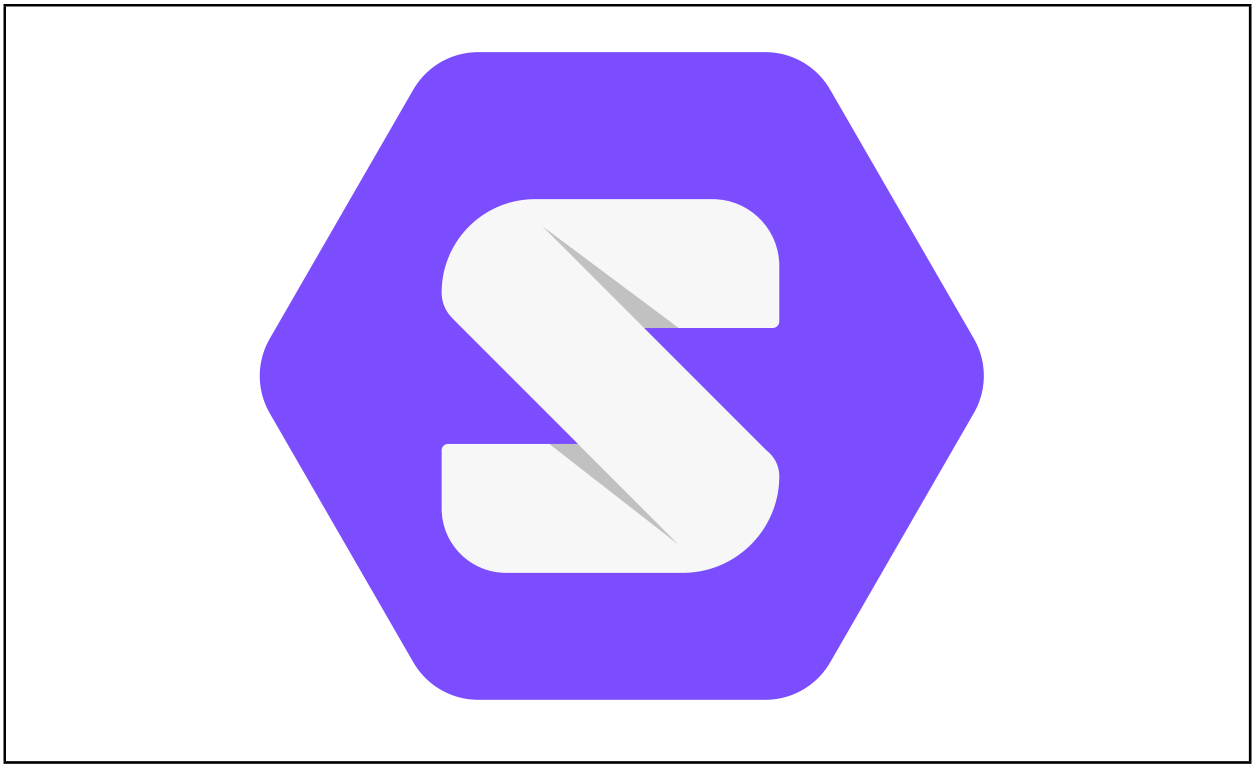
data market

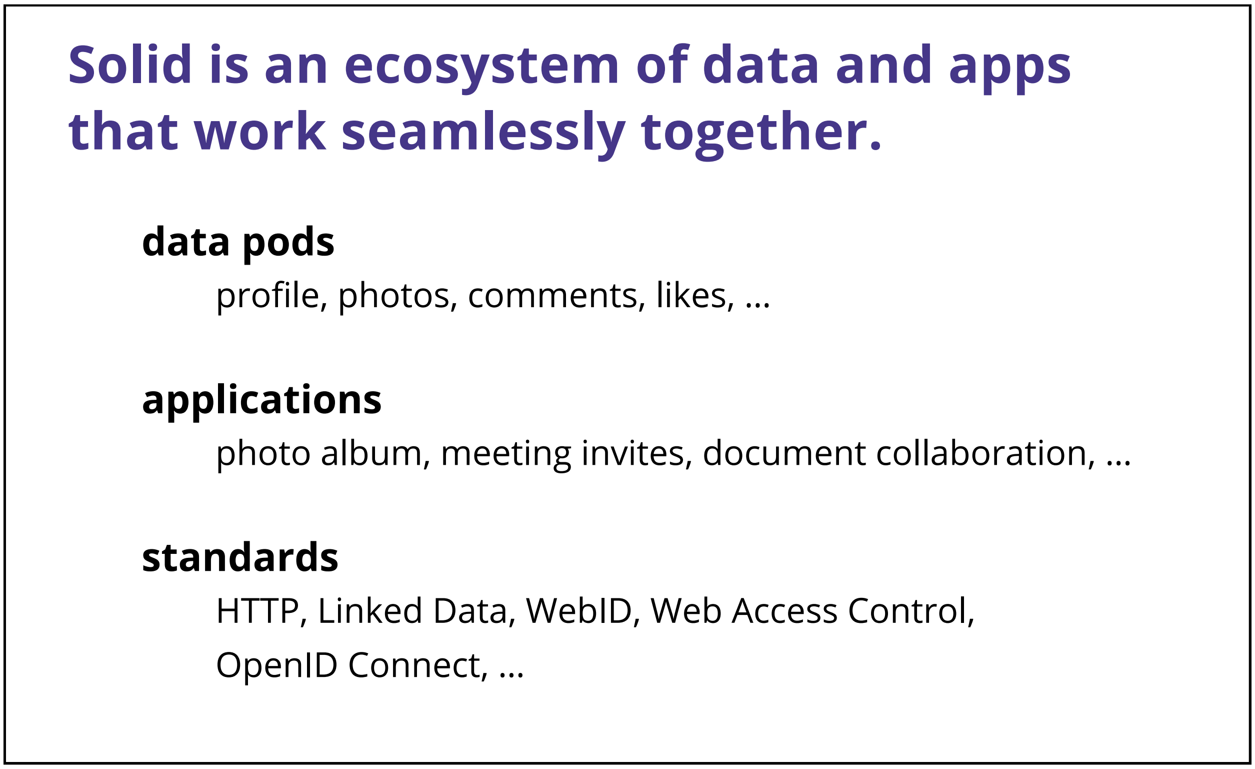
data pod
unlimited storage

competition based on service quality

data pod
secure backups

data pod
high-speed transfer





Solid is an ecosystem of data and apps that work seamlessly together.

data pods

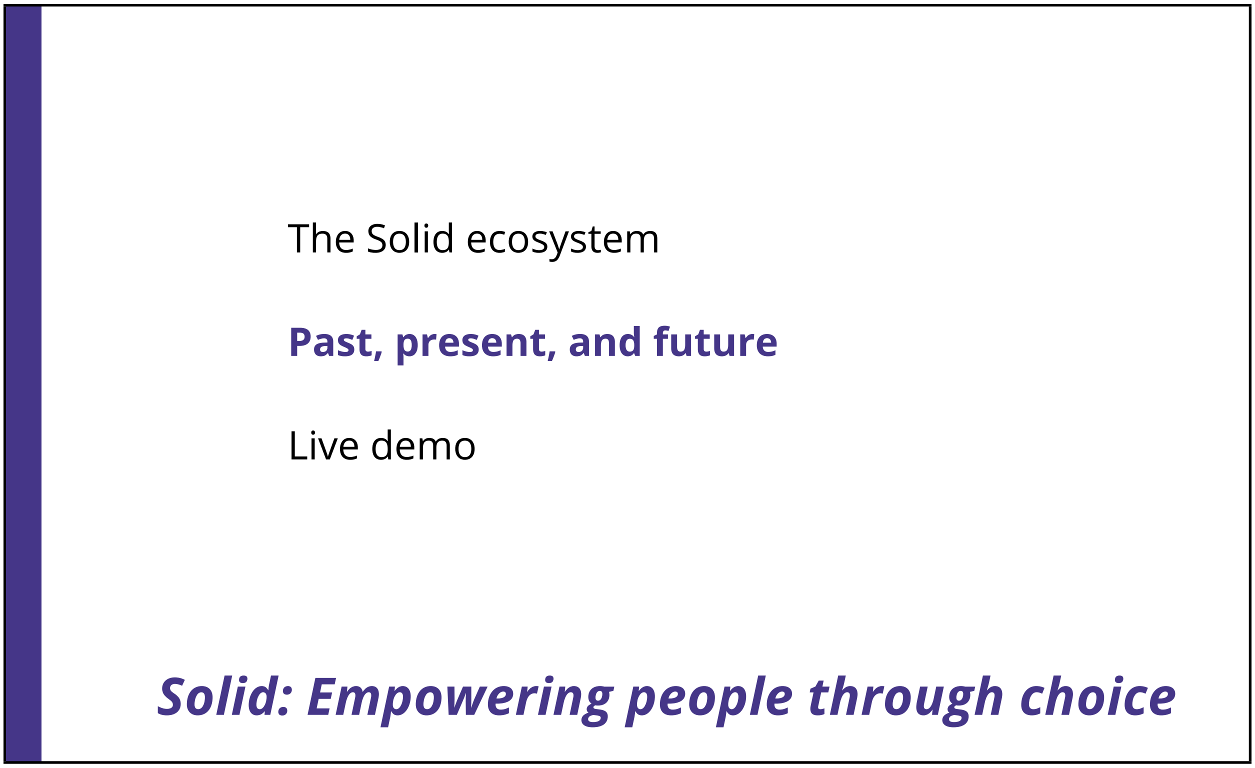
profile, photos, comments, likes, ...

applications

photo album, meeting invites, document collaboration, ...

standards

HTTP, Linked Data, WebID, Web Access Control, OpenID Connect, ...

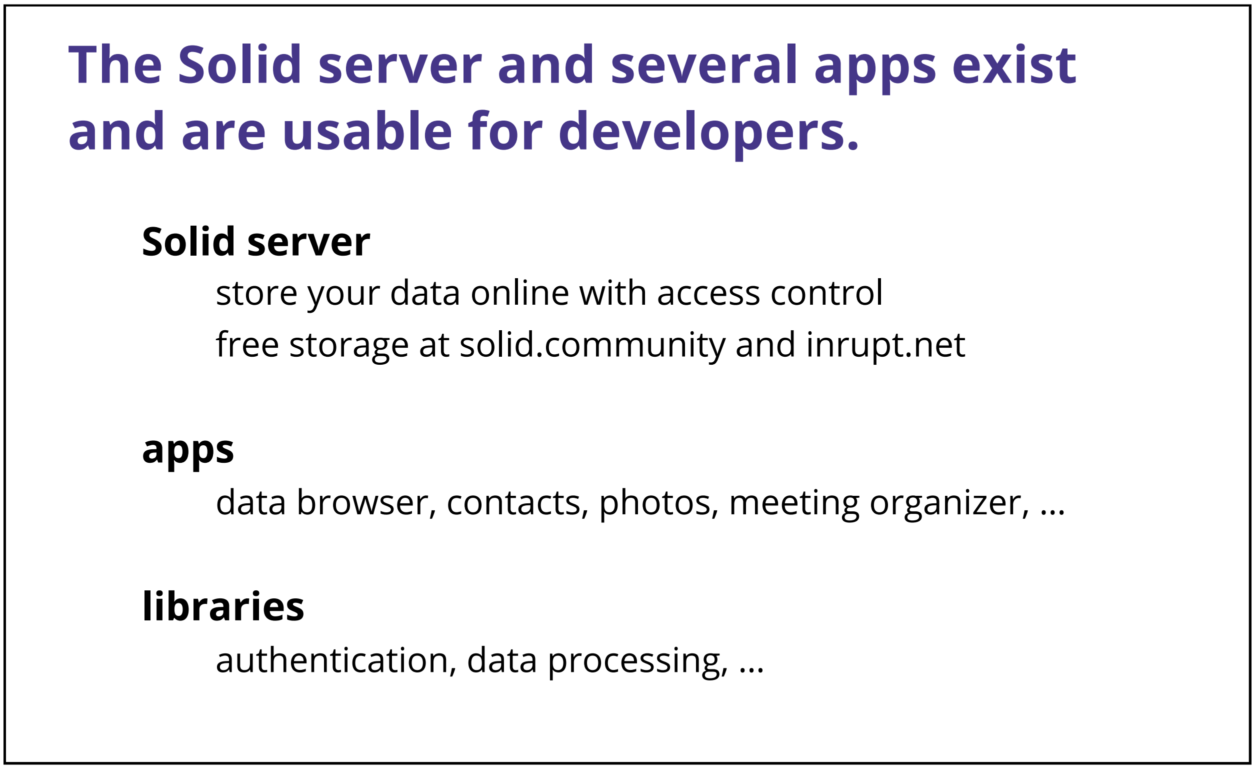


The Solid ecosystem

Past, present, and future

Live demo

Solid: Empowering people through choice



The Solid server and several apps exist and are usable for developers.

Solid server

store your data online with access control

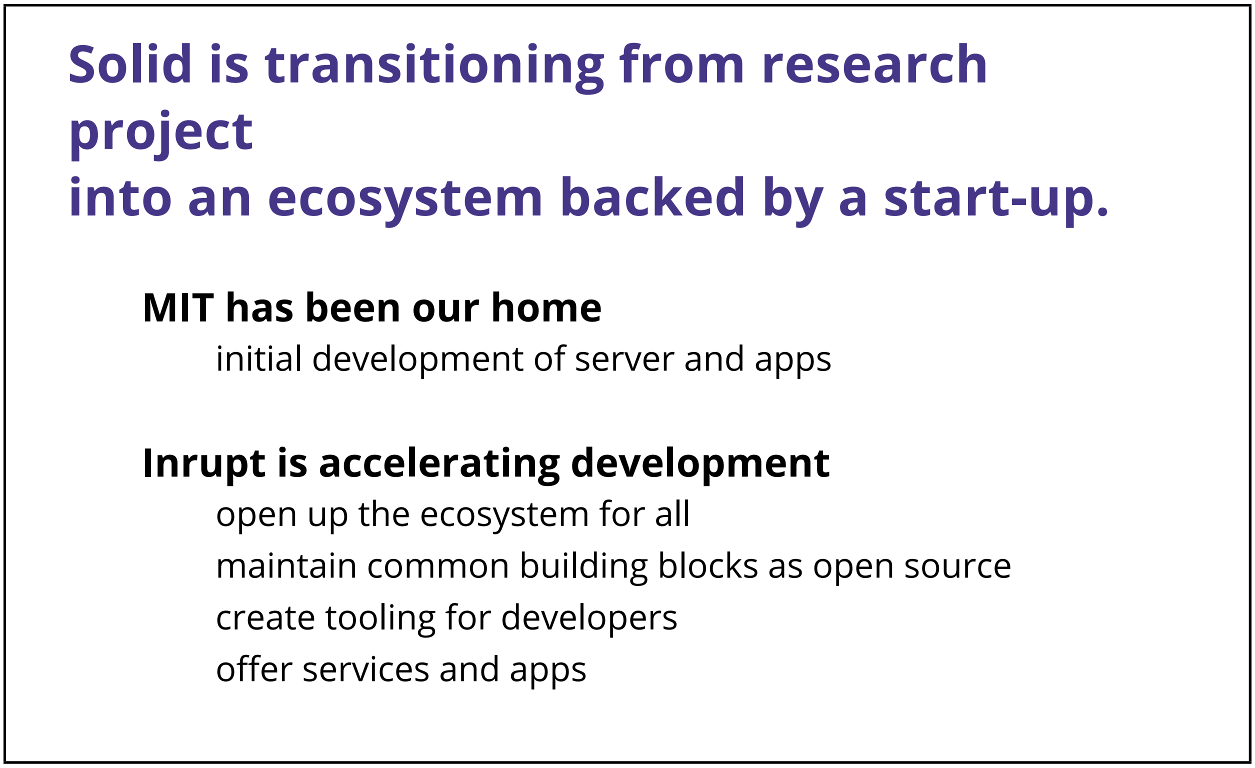
free storage at solid.community and inrupt.net

apps

data browser, contacts, photos, meeting organizer, …

libraries

authentication, data processing, …



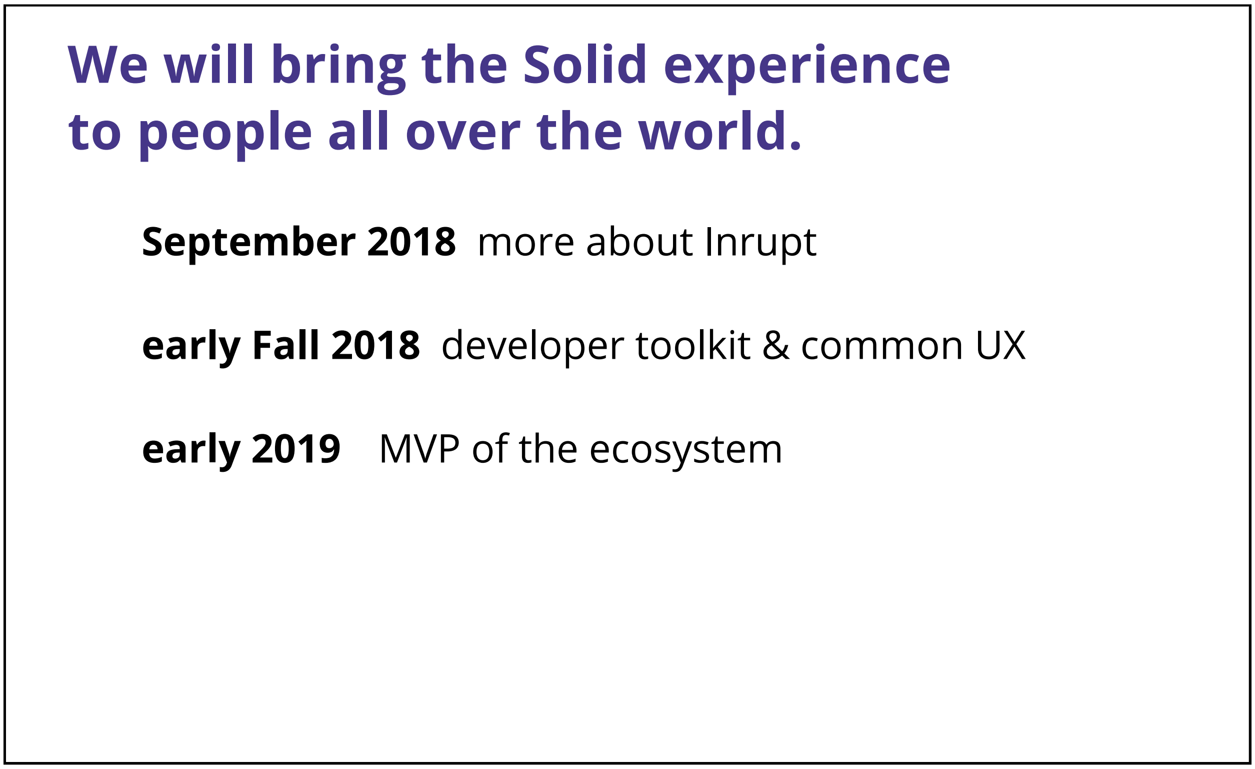
Solid is transitioning from research project
into an ecosystem backed by a start-up.

MIT has been our home

initial development of server and apps

Inrupt is accelerating development

open up the ecosystem for all
maintain common building blocks as open source
create tooling for developers
offer services and apps

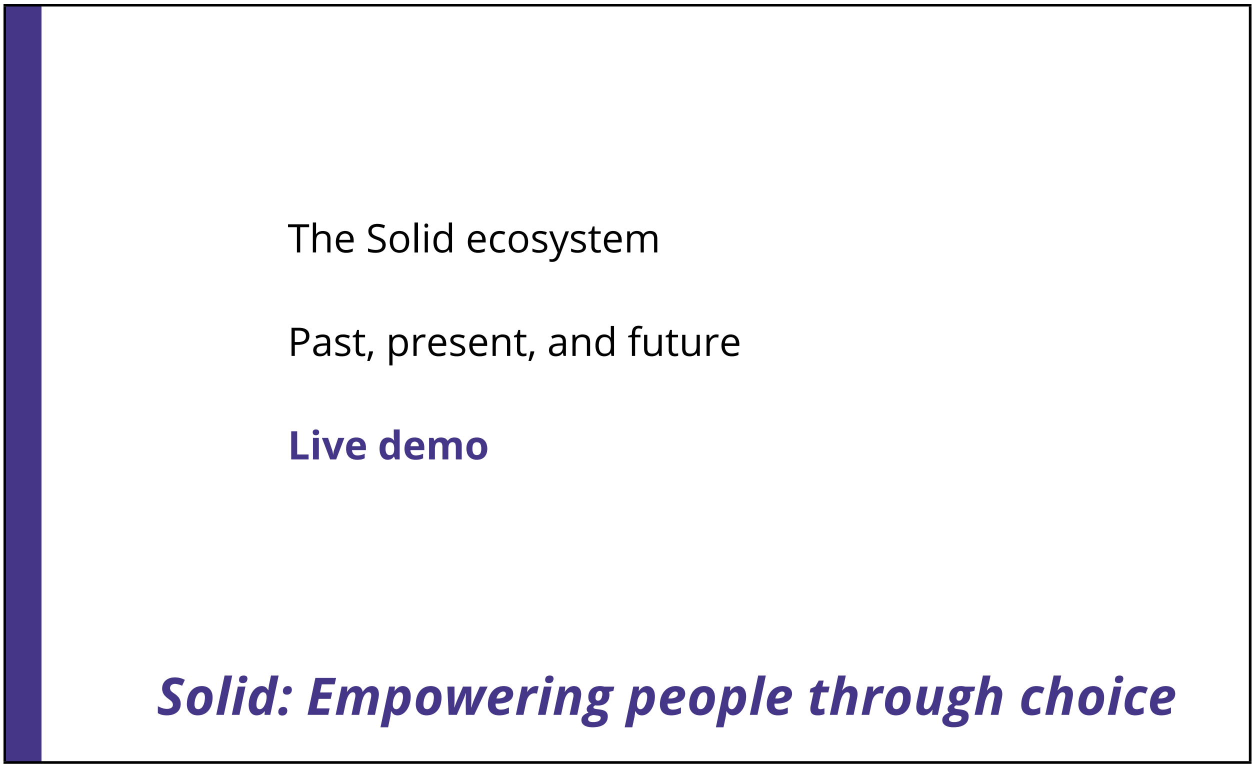


We will bring the Solid experience
to people all over the world.

September 2018 more about Inrupt

early Fall 2018 developer toolkit & common UX

early 2019 MVP of the ecosystem



The Solid ecosystem

Past, present, and future

Live demo

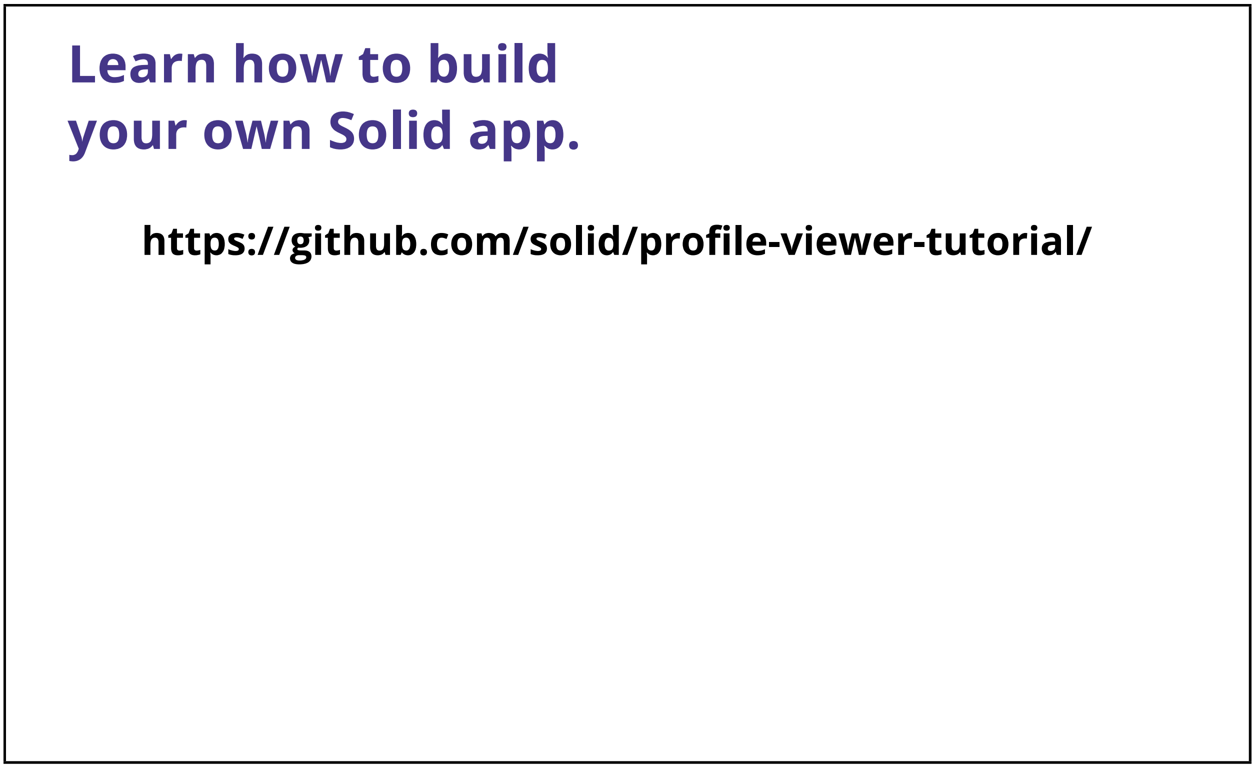
Solid: Empowering people through choice

Get your own Solid pod at one of these places.

https://solid.community/

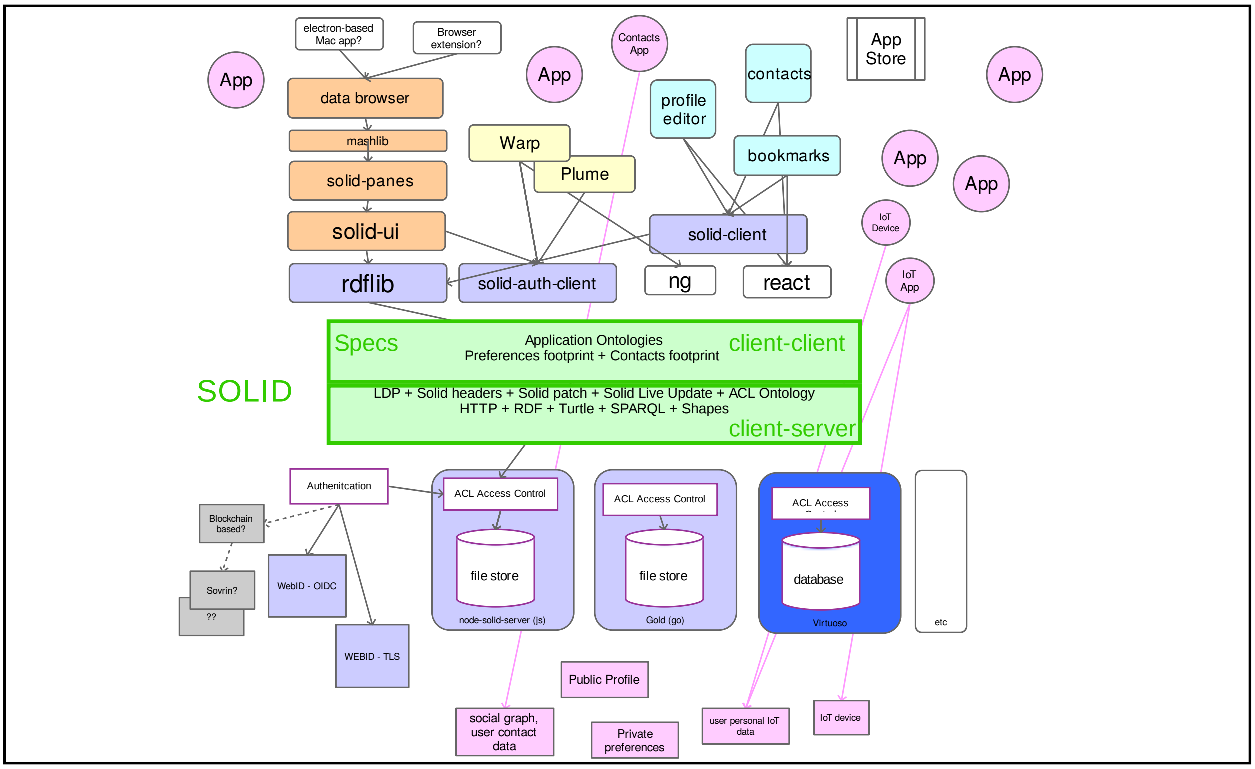
https://solidtest.space/

https://inrupt.net/



Learn how to build your own Solid app.

https://github.com/solid/profile-viewer-tutorial/



electron-based Mac app?

Browser extension?

Contacts App

App Store

App

data browser

mashlib

solid-panes

solid-ui

rdflib

Warp

Plume

solid-auth-client

profile editor

contacts

bookmarks

solid-client

ng

react

IoT Device

IoT App

SOLID

Specs

Application Ontologies
Preferences footprint + Contacts footprint

client-client

LDP + Solid headers + Solid patch + Solid Live Update + ACL Ontology
HTTP + RDF + Turtle + SPARQL + Shapes

client-server

Authenitcation

Blockchain based?

Sovrin?

??

WebID - OIDC

WEBID - TLS

ACL Access Control

file store

node-solid-server (js)

ACL Access Control

file store

Gold (go)

ACL Access

database

Virtuoso

etc

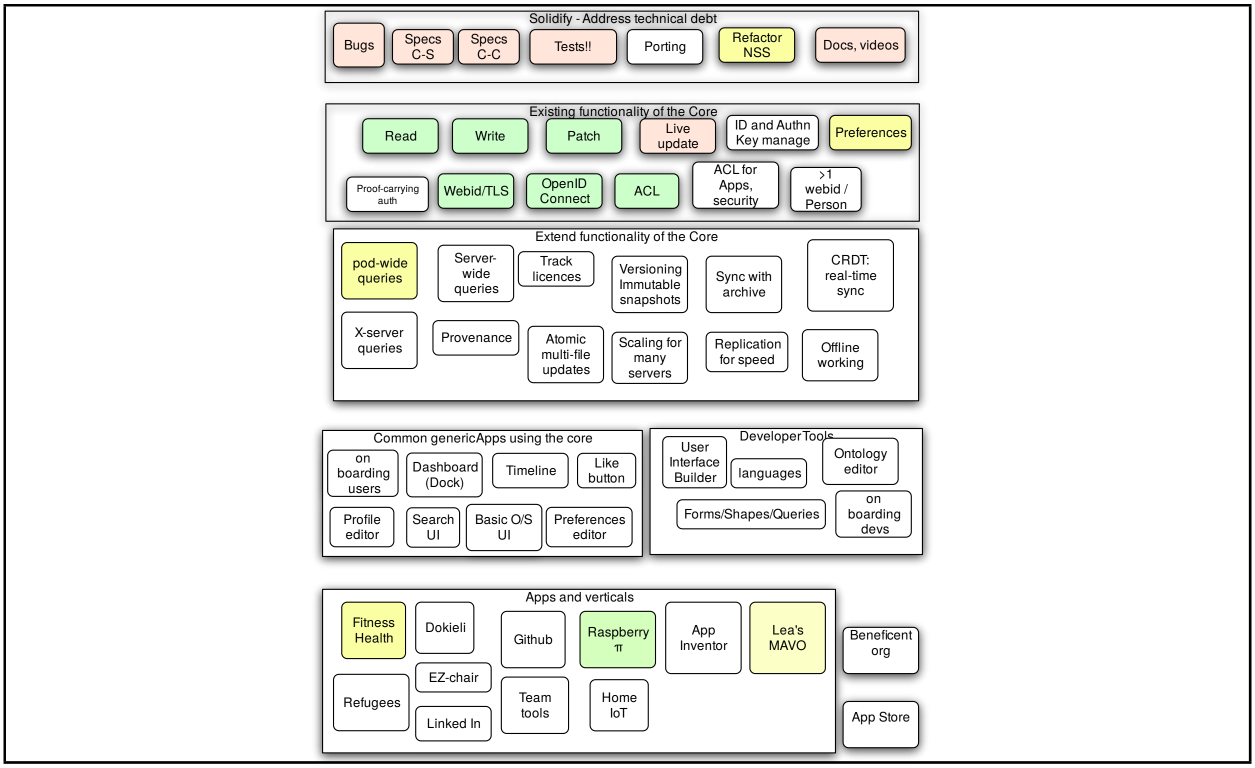
Public Profile

social graph, user contact data

Private preferences

user personal IoT data

IoT device



Solidify - Address technical debt

Bugs | Specs C-S | Specs C-C | Tests!! | Porting | Refactor NSS | Docs, videos

Existing functionality of the Core

Read | Write | Patch | Live update | ID and Authn Key manage | Preferences

Proof-carrying auth | Webid/TLS | OpenID Connect | ACL | ACL for Apps, security | >1 webid / Person

Extend functionality of the Core

pod-wide queries | Server-wide queries | Track licences | Versioning Immutable snapshots | Sync with archive | CRDT: real-time sync

X-server queries | Provenance | Atomic multi-file updates | Scaling for many servers | Replication for speed | Offline working

Common genericApps using the core

on boarding users | Dashboard (Dock) | Timeline | Like button

Profile editor | Search UI | Basic O/S UI | Preferences editor

DeveloperTools

User Interface Builder | languages | Ontology editor

Forms/Shapes/Queries | on boarding devs

Apps and verticals

Fitness Health | Dokieli | Github | Raspberry π | App Inventor | Lea's MAVO

Refugees | EZ-chair | Linked In | Team tools | Home IoT

Beneficent org

App Store