Database Protection: Resolving the U.S. Database Dilemma with an Eye toward International Protection

Michael Freno

Follow this and additional works at: http://scholarship.law.cornell.edu/cilj
Part of the Law Commons

Recommended Citation
Available at: http://scholarship.law.cornell.edu/cilj/vol34/iss1/5

This Note is brought to you for free and open access by Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell International Law Journal by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact jmp8@cornell.edu.
Database Protection: Resolving the U.S. Database Dilemma with an Eye Toward International Protection

Michael Freno*

Introduction .................................................... 166
I. Database Protection in the United States ................. 167
   A. Present Copyright Protection of Databases in the United States ...................................... 168
      2. Feist Publications, Inc. v. Rural Telephone Service Co.: The Originality Requirement .... 169
   B. Legislative Proposals for Protecting Non-creative Databases ........................................ 171
      1. Database Investment and Intellectual Property Antipiracy Act (House Bill 3531) .......... 171
      2. Collections of Information Antipiracy Act (House Bill 2652) .................................. 172
      3. Collections of Information Anti-Piracy Act (House Bill 354) .................................. 174
      4. Consumer and Investor Access to Information Act (House Bill 1858) ......................... 175

II. The International Framework for Database Protection .... 176
   A. International Copyright Treaties and Related Trade Agreements ..................................... 177
      1. Berne Convention ...................................... 177
      2. WIPO Copyright Treaty and Proposal .............. 178
      3. TRIPS ................................................. 180
   B. EU Database Directive .................................. 182

III. A Compromise Proposal for U.S. Database Protection .... 184
   A. An Argument for Additional Protection .......... 185
      1. The Argument for Protecting Non-creative Databases ........................................ 185

34 Cornell Int'l. L.J. 165 (2001)
Introduction
Imagine. Micromedex, Inc., a provider of in-depth medical information, invests hundreds of thousands of dollars developing a comprehensive and reliable database of poison antidotes called Poisindex. Micromedex expects to recoup its costs and make a profit by licensing the database to

1. This fictional example is adapted from Anne E. Kornblut, Database Compilers Fight for Copyright Protection, BOSTON GLOBE, Sept. 21, 1999, at A1.
hospitals and health clinics in the United States and abroad. The database provides an invaluable tool for the public, and Micromedex sells its first batch of copies in the United States. But then sales stop. Micromedex discovers that the database was copied onto the Internet and eventually copied in the United States and Europe. Consulting its attorneys, Micromedex discovers that the copying was perfectly legal in each instance, and there was virtually nothing it could have done to protect its database from piracy.\(^2\) Unable to sell another copy, Micromedex writes off the loss and halts all investment in databases. Moreover, observing Micromedex’s hard lesson, every other private investor avoids databases. As a result, U.S. businesses fail to create countless databases that could benefit the public. This example illustrates the database dilemma in the United States.

Over the past decade, Congress has contemplated several bills for protecting non-creative databases\(^3\) such as Micromedex’s Poisindex. The issue is which proposal the United States should adopt. This Note offers a proposal for legislation to resolve the database dilemma in both the United States and abroad. Part I describes the history of database protection in the United States and recent legislative proposals for resolving the dilemma, including strong bills based on creating a sui generis\(^4\) right in certain forms of data and weak bills based on misappropriation law. Part II discusses the current international framework for database protection with an emphasis on a regional agreement adopted by the European Union in 1996, which created a sui generis right for protecting non-creative databases. Part III argues that because the two current legislative proposals before Congress are flawed—one being too weak and the other too strong—a compromise proposal is required. The compromise attempts to strike a balance between creating an incentive for database makers\(^5\) to invest in databases, obtaining international protection for U.S. database makers, and preventing debilitating monopolies over data in the United States.

I. Database Protection in the United States

In the United States, copyright law provides scant protection for databases.\(^6\) In what follows, the present state of database protection in the

---

2. This is not entirely true. Although Micromedex could not protect its database under federal copyright law, it may have been able to protect itself from piracy through licensing agreements and other self-created protections. See infra Part III.

3. For a description of non-creative databases, see infra Part I.

4. The European Union Database Directive uses this term to describe the right created for the protection of non-creative databases. This term is appropriate because such a right creates a new regime of legal protection that finds no origin in copyright laws, misappropriation laws, or any other area of U.S. law.

5. The term “database maker” includes the creator of the database and the entity owning the database.

6. Discussion of U.S. copyright law covers federal copyright, not state copyright. State copyright “claims are either preempted by the Copyright Act or provide protection that is useless for mass-produced, readily-accessible databases.” Jeffrey C. Wolken, Note, Just The Facts, Ma’am. A Case For Uniform Federal Regulation of Information Databases in The New Information Age, 48 SYRACUSE L. REV. 1263, 1291 (1998).
United States is presented in two subsections. The first describes present protection of creative databases under the Copyright Act. The second describes recent legislative attempts to extend protection to non-creative databases.

A. Present Copyright Protection of Databases in the United States


Although creative databases are potentially copyrightable as "compilations" under the Copyright Act, an axiom of copyright law is that only expressions are protectable, not facts or ideas. A database of facts stands on dubious ground. Historically, courts extended copyright law to cover some factual compilations under the judicially created "sweat of the brow" doctrine. Under this doctrine, protection could be afforded to a factual compilation if the author showed sufficient effort and expense in generating the final product. Database protection under this regime bore the following rough characteristics:

| Source | The Federal Copyright Act combined with the common law "sweat of the brow" doctrine protected non-creative databases. |
| Scope | Copyright protection applied to (1) factual compilations that were (2) created at sufficient labor and expense to the compiler. |
| Rights | The owner of a copyright in a non-creative database had the exclusive right to reproduce the work. |
| Exceptions | The rights of copyright ownership were limited by the "fair use" defense and other important exceptions. |
| Duration | Copyright ownership continued for the life of the author, or compiler, plus fifty years. |

8. See id. § 101. A compilation is "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." Id.
9. Although the Copyright Act expressly excludes ideas as copyrightable subject matter, it does not expressly exclude the protection of facts. See 17 U.S.C. § 102(b). However, § 102(b) does exclude "discoveries" from protection, and the Supreme Court has held that facts are not a copyrightable subject matter. Harper & Row, Publishers, Inc. v. Nation Enters., 471 U.S. 539, 548 (1985). In Harper, the Court held that facts cannot be protected because they lack originality: "[C]opyright does not prevent subsequent users from copying from a prior author's work those constituent elements that are not original—for example . . . facts, or materials in the public domain—as long as such use does not unfairly appropriate the author's original contributions." Id.
10. See generally MARSHALL LEAFFER, UNDERSTANDING COPYRIGHT LAW 69-71 (3d ed. 1999) (explaining "sweat of the brow").
11. Id. at 69.
12. For an enumeration of the exclusive rights of a copyright owner, see 17 U.S.C. § 106.
13. 17 U.S.C. § 107. Fair use is a defense to copyright infringement that allows a third party to use a copyrighted work in a reasonable manner without the copyright owner's consent. LEAFFER, supra note 10, at 427-69.
14. 17 U.S.C. § 302(a). Duration is now seventy years. Also, the duration is based on certain assumptions—for example that the work was created after January 1, 1978. See LEAFFER, supra note 10, at 223-36.
For example, in *Jeweler's Circular Publishing Co. v. Keystone Publishing Co.*, the Second Circuit held that whether a book may be protected does not depend on whether the materials that make up the book consist of matters that "show literary skill or originality, either in thought or in language, or anything more than industrious collection."\(^{15}\) Thus, sufficient labor could inject a factual compilation into the realm of copyright law. Under this regime, Micromedex would have a chance to protect its Poisindex against piracy in the United States.

### 2. Feist Publications, Inc. v. Rural Telephone Service Co.: The Originality Requirement

In 1991, the Supreme Court eliminated the "sweat of the brow" doctrine in *Feist Publications, Inc. v. Rural Telephone Service Co.* by imposing an originality requirement on every copyrightable work.\(^{16}\) In *Feist*, a local phone company brought an unsuccessful suit against a publishing company for copying the white pages of its phone book.\(^{17}\) The Court held that an alphabetical listing of people and their addresses was not sufficiently "original."\(^{18}\) The Court emphasized that the primary objective for copyright under the Constitution is "to promote the Progress of Science and useful Arts," not to reward the labors of authors.\(^{19}\) As a result, the Court found that a database may receive copyright protection only if the selection, coordination, or arrangement of its facts are sufficiently original, where "originality" is understood to mean that the work was not copied and contained a modicum of creativity.\(^{20}\) As a result, *Feist* terminated all copyright protection for non-creative databases, regardless of the expense and labor incurred by the creators of databases.\(^{21}\)

### 3. Post-1991 Protection of Databases: Creative Databases

Although *Feist* leaves room for the protection of creative databases, this protection is thin. Most post-*Feist* litigation in the area of copyright infringement of databases involves the level and nature of creativity required to trigger copyright protection. The courts have developed a few methods for measuring the requisite creativity.\(^{22}\) One such test is whether

---

15. 281 F. 83, 88 (2d Cir. 1922) (emphasis added).
17. Before copying the pages, the publishing company first asked the telephone company to license their use of the directory, but the telephone company was refused. *Id.* at 343.
18. *Id.* at 361.
19. *Id.* at 349. The U.S. Constitution states that Congress shall have the 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." *U.S. Const.* art. I, § 8.
20. This language is also now part of the definition of "compilation" in the Copyright Act. 17 U.S.C. § 101 (1999).
the selection of the data constitutes "thoughtful" selection. For example, the plaintiff in *Key Publications, Inc. v. Chinatown Today Publishing Enterprises, Inc.* collected business cards, copied restaurant listings from another directory, and ultimately produced a "yellow pages" for New York's Chinese business community, which included about 9000 listings organized into 260 business categories. The defendant allegedly copied 1500 of the plaintiff's listings. The Second Circuit concluded that the "database" was sufficiently creative to receive copyright protection.

Another guide for measuring creativity in databases is to look at how its creators selected and arranged the data as a whole. In the highly publicized case of *Warren Publishing, Inc. v. Microdos Data Corp.*, Warren Publishing sued Microdos for allegedly copying its factbook containing information about all cable television systems in the United States and selling an electronic version of it. The Eleventh Circuit held that Warren Publishing could not protect its factbook under copyright law because the book lacked sufficient creativity. The court focused on how Microdos selected and arranged the data as a whole and concluded that Microdos copied no original selection, coordination, or arrangement of Warren's factual compilation. According to the court, copying data is not considered infringement if the selection and arrangement in the new work is not substantially similar to the selection and arrangement in the original.

A third method for determining whether a database contains adequate creativity is the "merger doctrine"—where there is only one or a few ways of expressing an idea, the expression merges with the idea and cannot be copyrighted. It follows from this doctrine that functional selections or arrangements of data according to basic organizing principles such as alphabetizing, cross-referencing, geographic areas, or chronological order are not protected.

---

24. *Key Publications*, 945 F.2d at 511.
25. *Id.* at 513. However, the court ultimately held for the defendant for unrelated reasons. One commentator described the element of creativity:

[The database] cited her initial judgment to select only those New York businesses that would be of interest to the Chinese-American community, and her testimony that she subsequently exercised discretion in excluding individual businesses that she thought would not remain open for very long. These acts of "thought and creativity" distinguished the Key directory from the typical white-pages, whose compiler "'slavishly' includes every available listing.


26. 115 F.3d 1509, 1510-14 (11th Cir. 1997).
27. *Id.* at 1520-21.
28. *Id.*
29. LEAFFER, *supra* note 10, at 82.
30. This doctrine precludes from protection compilations that classify churches by denomination, attorneys by areas of specialty, cable systems by principal community served, and radiators by manufacturer and application. Richard L. Stone & John D. Pernick, *Protecting Databases: Copyright? We Don't Need No Stinkin' Copyright*, 16 COMPUTER LAW 17 (1999).
Each method for measuring creativity in databases suggests that it is difficult to meet the standard of creativity in the ordinary, useful database. Indeed, the more comprehensive and useful the selection or arrangement of the data, the less likely it will be protected by copyright.31

The following table summarizes the present protection afforded to databases in the United States under copyright law:

<table>
<thead>
<tr>
<th>Source</th>
<th>The Federal Copyright Act protects some creative databases.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>Copyright protection applies to (1) factual compilations that are (2) selected, coordinated, or arranged in a sufficiently creative way.</td>
</tr>
<tr>
<td>Rights</td>
<td>The owner of a copyright in a creative database has the exclusive right to reproduce the work, create derivative works based on the database, and other related rights.32</td>
</tr>
<tr>
<td>Exceptions</td>
<td>The rights of copyright ownership are limited by the &quot;fair use&quot; defense33 and other important exceptions.</td>
</tr>
<tr>
<td>Duration</td>
<td>Copyright ownership continues for the life of the author plus seventy years.34</td>
</tr>
</tbody>
</table>

Assuming that Micromedex's database lacks the required creativity by presenting the poison antidotes in alphabetical order and categorizing them according to industry standard, then present copyright law would provide no protection in the United States.

B. Legislative Proposals for Protecting Non-creative Databases

Since 1996, Congress has considered several bills designed to deal with the database dilemma. Although each proposal attempts to replace the originality requirement with some form of the "sweat of the brow" doctrine, they differ significantly on the level of protection afforded to non-creative databases. Some proposals would provide a high level of protection based on a sui generis property right in non-creative databases, a regime modeled on the European Union's Database Directive.35 Others would provide a low level of protection modeled on misappropriation doctrines and would prevent pirates from extracting data from non-creative databases in certain contexts. However, no proposal has survived the Senate.

1. Database Investment and Intellectual Property Antipiracy Act (House Bill 3531)

The 104th Congress considered the Database Investment and Intellectual Property Antipiracy Act,36 which went outside copyright law and sought to create a new and unique, or sui generis, right protecting certain

---

31. See id.
34. 17 U.S.C. § 302(a). This duration is based on certain assumptions—for example, the work was created after January 1, 1978. See LEAFFER, supra note 10, at 223-36.
35. See supra Part II.B.
The basic contours of the proposal included the following:

<table>
<thead>
<tr>
<th>Source</th>
<th>The proposed statute would have created a sui generis right outside copyright law that protected non-creative databases.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>Protection would have applied to (1) collections or compilations (in any medium) arranged in a systematic or methodical way, (2) that are the result of a qualitatively or quantitatively substantial investment of human, technical, financial, or other resources in the collection, assembly, verification, organization or presentation of the database contents if (3) the database is used, reused, or intended for use or reuse in commerce.</td>
</tr>
<tr>
<td>Rights</td>
<td>The statute would have prohibited (1) the extraction of all or a substantial part of the contents of a database in a way that (2) conflicts with the database owner's normal exploitation of the database or adversely affects the actual or potential market for the database.</td>
</tr>
<tr>
<td>Exceptions</td>
<td>The statute would have provided no exception for personal, educational, scientific, or research uses.</td>
</tr>
<tr>
<td>Duration</td>
<td>The property right in the database would continue for a twenty-five-year period, but this could restart upon any significant change to the database.</td>
</tr>
<tr>
<td>Remedies</td>
<td>The statute would have provided extensive civil—-injunctive relief, impoundment, and monetary relief—and criminal penalties for a violation.</td>
</tr>
</tbody>
</table>

This proposal was broad in scope because it did not require even a modicum of creativity, required a relatively small amount of extraction to trigger infringement, gave few exceptions, and left the potential for perpetual protection through ongoing changes to the database. For these and other reasons, the proposal died in the Judiciary Committee.

2. Collections of Information Antipiracy Act (House Bill 2652)

The successor bill in the 105th Congress was the Collections of Information Antipiracy Act (House Bill 2652). This proposal was modeled on the EU Database Directive. See supra Part II.B. Under the proposal, “database" means a collection, assembly or compilation, in any form or medium now or later known or developed, of works, data or other materials, arranged in a systematic or methodical way. The term of protection would have been twenty-five years from January 1 after the date that the database is placed into commercial use or made available to the public, whichever is earlier. Presumably, the proposal would cover phone books, history or math books, literary works, and musical works. Under the proposal, “extraction" means the permanent or temporary transfer of all or a substantial part of the contents of a database or of a copy or copies thereof. The bill did “not provide an exception for personal or scientific uses."
tion Antipiracy Act. 48 Like its predecessor, this bill also would have created a sui generis right; however, it attempted to protect non-creative databases by adding a chapter titled “Misappropriation of Collections of Information” to the Copyright Act. 49 This proposal comprised the following features.

The Collections of Information Antipiracy Act: House Bill 2652

| Source | The proposal would have created a sui generis right protecting non-creative databases by expanding the Copyright Act with a chapter entitled “Misappropriation of Collections of Information.” |
| Scope | Protection would have applied to (1) collections of “facts, data, works of authorship, or any other intangible material capable of being collected and organized in a systematic way” 50 that were (2) gathered, organized, or maintained through the investment of substantial monetary or other resources and that were (3) offered or intended to be offered for sale or otherwise in commerce. |
| Rights | The statute would have prohibited (1) the extraction of all or a substantial part of the contents of the database that (2) caused harm to the actual or potential market of the owner. 51 |
| Exceptions | The bill enumerated a number of exceptions including (1) extraction or use of insubstantial parts of a protected collection (e.g., individual items of information), (2) use of protected information to verify independently gathered data, (3) non-profit research in a manner that does not hurt the actual or potential market for the database, (4) reporting the news, and (5) gathering information by other means. 52 Moreover, the bill excluded government collections from protection. |
| Duration | The bill specified no duration for protection. 53 |
| Remedies | The bill prescribed civil remedies—injunctive relief; impoundment; and monetary relief, including damages and attorneys' fees, 54—and criminal penalties, including fines ranging from $250,000 to $500,000 and imprisonment for five to ten years. 55 |

Although House Bill 2652 passed in the House of Representatives twice and accommodated the criticisms aimed at its predecessor, the Sen-

---

49. H.R. 2652, § 3.
50. Id. (proposing 17 U.S.C. § 1201(2)).
52. Id. (proposing 17 U.S.C. § 1204).
53. Conley et al., supra note 23, ¶ 88 (footnote omitted).
54. H.R. 2652, § 3 (proposing 17 U.S.C. § 1206). Injunctive relief and impoundment would not be available against the federal government. Id.
ate rejected it. To explain the Senate's disapproval, one commentator focused on three problems with the proposal: (i) its broad scope, (ii) its lack of a sufficient exception for educational and research uses of otherwise protected data, and (iii) its potential for perpetual protection.

First, the bill contained the term "collection of information," defining "information" as "facts, data, works of authorship, or any other intangible material capable of being collected and organized in a systematic way." The problem was that almost "anything could be defined as a collection of information, and since there is no time limit to the protection, there will be perpetual protection beyond the recoupment of investment. Such an idea conflicts with the values of a free market economy." Second, despite an apparent exception from liability for some educational and research uses, the bill's specific exception for not-for-profit scientific or research uses did not apply where the use "harm[s] the actual or potential market for the product." Thus, provisions of the bill render useless the immunity for educational, scientific, and research uses where such uses cause harm, and arguably, all unauthorized uses cause actual or potential harm to the database owner. Third, the failure to specify a duration for protection creates the possibility of perpetual privileges.

3. Collections of Information Anti-Piracy Act (House Bill 354)

The 106th Congress considered the Collections of Information Anti-Piracy Act, which, like House Bill 2652, would amend the Copyright Act by adding a new chapter. The proposal contained the following characteristics:

56. Julius J. Marke, Database Protection Bills Pending in Congress, N.Y.L.J., Aug. 17, 1999, at 5. "The House passed H.R. 2652 in May 1998, and then passed it again as part of the House version of what was to become the Digital Millenium [sic] Copyright Act. However, the Senate deleted the provisions of H.R. 2652 from the DMCA before both houses passed that legislation in October 1998." Conley et al., supra note 23, ¶ 90.


58. H.R. 2652, § 3 (proposing 17 U.S.C. § 1201(2)).

59. Tessensohn, supra note 57, at 471. In other words, the broad scope combined with the unlimited duration creates too much protection unnecessarily.


61. Tessensohn, supra note 57, at 472.

62. Id. at 473.

The Collections of Information Anti-Piracy Act: House Bill 354

| Source | The proposal would create a sui generis right protecting non-creative databases by expanding the Copyright Act with a Chapter called “Misappropriation of Collections of Information.”
| Scope | Protection would apply to (1) factual databases that were (2) created and maintained by a substantial monetary investment or other resources and (3) used in commerce.
| Rights | The statute would prohibit (1) the extraction of all or a substantial part of the contents of the database that (2) caused harm to the actual or potential market of the owner.
| Exceptions | The statute would offer lots of exceptions and exclusions such as fair use measures for scientific, education, and research purposes; government collections; and news reporting.
| Duration | The property right in the database would continue for a fifteen-year period beginning when the database was injected into commerce, but could restart upon any significant change to the database.
| Remedies | The statute would prescribe civil remedies—injunctive relief; impoundment; and monetary relief, including damages and attorneys’ fees—and criminal penalties, including fines ranging from $250,000 to $500,000 and imprisonment for five to ten years.

Although this bill was nearly identical to its predecessor, it included a fair use exception and provided a fifteen-year limit to protection. Under this regime, Micromedex could protect its Poisindex database from piracy if the pirate copied and exploited a substantial portion of the database’s factual content. Congress did not take action on this proposal.

4. Consumer and Investor Access to Information Act (H.R. 1858)

Legislators also considered the Consumer and Investor Access to Information Act. This proposal “features a relatively narrow prohibition that is aimed mainly at preventing wholesale misappropriation of databases, prohibiting duplication that produces a database that is ‘substantially the same’ as the first.” The bill may be summarized as follows:

64. Id. § 2.
66. Id. (proposing 17 U.S.C. §§ 1403-1404). See infra Part III for further discussion of these exceptions.
68. H.R. 354 § 2 (proposing 17 U.S.C. § 1406). Injunctive relief and impoundment are not available against the federal government. Id.
70. See supra note 13.
The Consumer and Investor Access to Information Act: House Bill 1858

<table>
<thead>
<tr>
<th>Source</th>
<th>The proposed statute would protect some databases from misappropriation.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope Protection</td>
<td>would apply to (1) collections of facts collected and organized in a</td>
</tr>
<tr>
<td></td>
<td>single place, (2) through the investment of substantial monetary or</td>
</tr>
<tr>
<td></td>
<td>other resources, and (3) for the purpose of providing access to those</td>
</tr>
<tr>
<td></td>
<td>discrete items of information by users of the database.73</td>
</tr>
<tr>
<td>Rights</td>
<td>The statute would prohibit the extraction of information from a database:(1) to create a new database that is substantially the same as the original database, (2) to sell or distribute to the public the new database, or (3) to sell or distribute in commerce the new database in competition with the original database.74</td>
</tr>
<tr>
<td>Exceptions</td>
<td>The statute would have numerous exceptions and exclusions such as collecting or using information obtained through other sources; news reporting; law enforcement and intelligence activities; scientific, educational, or research uses;75 government information; databases related to Internet communications; computer programs; and individual facts, ideas, and related concepts.76</td>
</tr>
<tr>
<td>Duration</td>
<td>This form of protection has no durational limit.</td>
</tr>
<tr>
<td>Remedies</td>
<td>Only the Federal Trade Commission has jurisdiction to bring an action against an entity for violating the statute. The statute mentions no remedies except in the context of securities market information.77</td>
</tr>
</tbody>
</table>

Unlike the other proposals, House Bill 1858 only protected the database if the pirate creates a new database that is substantially the same as the old one. Thus, under this regime, Micromedex could only protect its database from wholesale copying. Moreover, this proposal offered little chance for increasing the protection of U.S. databases from international piracy.78

II. The International Framework for Database Protection

Strictly speaking, there is no international copyright law.79 Rather, the international framework for copyright protection consists of the national

---

73. H.R. 1858, § 101.
74. Id. § 102.
75. Id. § 103(d).
76. Id. §§ 103-104.
77. Id. §§ 107, 201.
78. See infra Part III.B.
laws between countries, which are linked through formal treaties, trade agreements, and regional agreements. Generally, this framework provides no more protection for non-creative databases than present copyright law in the United States. Although the European Union adopted a regional agreement in 1996 that created a regime for protecting non-creative databases by engendering a sui generis right in such databases, no international instrument protects non-creative databases made in the United States.

A. International Copyright Treaties and Related Trade Agreements

1. Berne Convention

Chief among treaties linking national copyright laws is the Berne Convention for the Protection of Literary and Artistic Works (Berne Convention or Berne). The Berne Convention dates back to 1886 and currently consists of more than 130 member countries. The World Intellectual Property Organization (WIPO), a specialized agency of the United Nations, now administers the Berne Convention.

Historically, the Berne Convention played a major role in harmonizing national copyright laws by requiring that each member recognize certain minimum standards of protection. However, there are at least three problems with using the Berne Convention to protect non-creative databases.

First, although the Berne Convention protects collections, it fails to

80. Id.
81. See infra Part II.B.
83. REES & CHARLTON, supra note 82 at 129.
85. In particular, the Berne Convention requires that member countries provide: (1) copyright owners and authors with a set of minimum rights; (2) restrictions on the scope of permissible limitations and exceptions; (3) a minimum term of protection for the life of the author plus fifty years; (4) a prohibition on formalities as a condition for enjoying rights; (5) retroactive protection for existing works in which copyright has not expired; and (6) "national treatment." National treatment prohibits a country from providing less favorable treatment to foreigners than to its own citizens with respect to intellectual property laws. Perlmutter, supra note 79.
86. Article 2(5) provides protection for collections of works to the extent that the selection and arrangement of their contents constitute intellectual creations. "It is argua-
protect both creative and non-creative databases. Berne seeks to protect and promote the international rights of authors, but it expressly excludes from protection "news of the day [and] . . . miscellaneous facts having the character of mere items of press information." In short, it provides no protection for U.S. databases abroad.

Second, even if the Berne Convention included protection for databases, a U.S. database maker could only expect protection equivalent to the Berne Convention's minimum standard, regardless of whether the United States provided a much higher level of protection. National treatment does not entail reciprocal treatment. Moreover, countries with few databases are likely to adopt the absolute minimum amount of protection, which may be insufficient for a U.S. citizen to enforce her rights under U.S. law in those countries.

Third, if the Berne Convention included a high level of protection for databases, a victim of piracy still can only enforce her rights outside her own country through an action in the foreign country's courts. In other words, the Berne Convention provides no impartial forum for addressing international copyright complaints. Although a Berne Convention member is obligated under the treaty to apply its own law to foreigners, a violation of the Berne Convention carries no consequences.

Although the Berne Convention offers no protection of databases at present, some commentators suggest ways in which the Berne Convention could be amended to protect non-creative databases. Today, any modification of the Berne Convention, or any other treaty for increasing database protection, would occur through WIPO.

2. WIPO Copyright Treaty and Proposal

In 1996, members of WIPO adopted the WIPO Copyright Treaty (WIPO Treaty), which "introduced new international rules to clarify the interpretation of existing intellectual property rights" without derogating any provision in the Berne Convention. Significantly, the WIPO Treaty specifically addresses compilations of data and provides the following regime for database protection.

ble whether 'protection for collections of works' can be interpreted to include protection of compilations of facts or data." Wayman, supra note 84, at 446 n.151.

87. Berne Convention, supra note 82, art. 2(8).

88. This protection derives from the minimum standards set by Berne. Id. arts. 5, 11, 12, 14.

89. National treatment may create a disincentive for some countries to provide strong national laws protecting databases. Indeed, a country would only want to protect its databases to the extent necessary to protect it from copying by its own citizens.

90. REES & CHARLTON, supra note 82 at 139.

91. Perlmutter, supra note 79, at 372-73.

92. See Wayman, supra note 84, at 446.


94. REES & CHARLTON, supra note 82, at 138.

95. WIPO Treaty, supra note 93, art. I, § 2
Protection of Databases under the WIPO Treaty of 1996: Present

Source
The WIPO Treaty is an international intellectual property treaty endorsed by roughly 160 countries.96

Scope
Protection applies to compilations of data, which, by reason of the selection or arrangement of their contents, constitute intellectual creations.97

Rights
In addition to having the moral rights of the author, the database author has the exclusive right to reproduce, translate, adapt, and broadcast the database.

Exceptions
The rights under the WIPO Treaty are limited by numerous exceptions, including: daily news, quotations from published works, allowance of exceptions for government information, public lectures, and educational purposes.

Duration
Protection continues for the life of the author plus fifty years.

Treatment
The WIPO Treaty requires national treatment among members.

Enforcement
A victim of piracy may only enforce her rights outside her own country through an action in the foreign country's courts.

Remedies
A victim of piracy may only seize infringing copies in the member country where the infringing copies are distributed or sold.

Because the WIPO Treaty requires creativity, it affords no greater protection to databases than the current U.S. Copyright Act. Moreover, although the WIPO Treaty theoretically protects creative databases, it suffers the same problems as the Berne Convention. Since the WIPO Treaty merely requires protection of creative databases, a U.S. database maker could only expect this low level of protection from a foreign Treaty member regardless of whether the United States provided a much higher level of protection. In addition, even if a U.S. citizen manufactured a database falling within the scope of protection afforded by a WIPO member, she could only enforce her rights in the courts of the country where infringement occurred. Moreover, the WIPO Treaty provides no penalties if the country ignores her rights.

Although WIPO contemplated adding protection of non-creative databases—in 1996 and 1997—through an amendment to the Berne Convention or through an entirely new instrument,98 lingering problems with enforcing rights and obtaining sufficient remedies cast doubt over whether the changes could give U.S. database makers sufficient international protection against piracy. The Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS)99 solves some of these concerns.

96. There are nearly 160 member nations of WIPO, including the United States and all fifteen Member States of the European Union. Wayman, supra note 84, at 446.
97. WIPO Treaty, supra note 93, art. V.
98. In 1996, WIPO hosted a diplomatic conference to consider several proposals including the Basic Proposal for the Substantive Provision of the Treaty on Intellectual Property in Respect of Databases to be Considered by the Diplomatic Conference (WIPO Proposal), Basic Proposal for the Substantive Provision of the Treaty on Intellectual Property in Respect of Databases to be Considered by the Diplomatic Conference, WIPO Doc. CRNR/DC/6 (Aug. 30, 1996). The conference did not adopt the proposal, but it adopted a resolution to continue work toward a treaty on database protection. Recommendation Concerning Databases, WIPO Doc. CRNR/DC/100 (Dec. 23, 1996); REES & CHARLTON, supra note 82, at 138.
3. TRIPS

TRIPS is the most important of several trade agreements that now contain copyright provisions. Visionaries in the area of intellectual property recognized that the General Agreement on Tariff and Trade (GATT) could help curb intellectual property piracy and force further uniformity among national intellectual property laws because GATT can impose trade sanctions on delinquent countries. As a result of the Uruguay round of GATT negotiations, which began in 1985 and ended in 1994, TRIPS was adopted in 1995. During the same negotiations, the World Trade Organization (WTO) was established to administer GATT, including TRIPS. TRIPS is now poised to revolutionize international intellectual property rights. Whereas the Berne Convention and the WIPO Treaty afforded isolated and essentially voluntary international legal protection, nations may now protect intellectual property through trade sanctions against infringing countries.

Generally, TRIPS incorporates the substantive norms of the Berne Convention by direct reference. In other words, TRIPS requires that its members provide minimum standards of intellectual property protection through their respective national laws. Thus, like the Berne Convention, TRIPS includes: (1) a set of minimum rights; (2) restrictions on the extent to which a member may limit protections; (3) a minimum term of protection for the life of the author plus fifty years; (4) a prohibition on formalities as a condition for enjoying rights; (5) retroactive protection for existing works in which copyright has not expired; and (6) a national treatment provision.

However, TRIPS differs from the Berne Convention in that it: (i) augments the scope of protection; (ii) requires most favored nation treat-
Protection of Databases under the TRIPS Agreement: Present

| Source | TRIPS is an international trade agreement containing intellectual property provisions. |
| Scope | Protection applies to (1) compilations of data (2) that by reason of the selection or arrangement of their contents constitute intellectual creations. |
| Rights | The database author has the exclusive right to reproduce, translate, adapt, and broadcast the database. |
| Exceptions | The rights are limited by numerous exceptions including: daily news, quotations from published works, allowance of exceptions for government information, public lectures, and uses for educational purposes. |
| Duration | Protection continues for the life of the author plus fifty years. |
| Treatment | The agreement provides national treatment and most favored nation treatment. |
| Enforcement | The agreement requires that each WTO member provide a detailed enforcement mechanism. Moreover, a country may enforce its rights against another country through the WTO dispute resolution system. |
| Remedies | A violation of TRIPS could open to the offending country to trade sanctions. |

Although TRIPS fails to protect non-creative databases, such as Micromedex's Poisindex, it offers several benefits over both the Berne Convention and the WIPO Treaty.

First, TRIPS may leave room for the protection of non-creative
Article 10(2) states that even though protection does not extend to the actual data contained in the compilations, it is "without prejudice to any copyright subsisting in the data or material" that a member state may choose to grant protection. Thus, if France, for example, protects non-creative databases under its national copyright law and someone in France copies Micromedex's Poisindex, then Micromedex may enforce the rights specified by French law via TRIPS.

Second, TRIPS supplies a possible mechanism by which a victim of piracy can enforce her rights. This avoids one of the major shortcomings of the Berne Convention and the WIPO Treaty; if the infringing country fails to honor its TRIPS obligations, it could face costly trade sanctions. Thus, if Micromedex's Poisindex fell within the scope of protection under TRIPS and France, a WTO member, failed to protect Micromedex's rights, the United States would have a specific mechanism by which to force France to comply with TRIPS.

Third, TRIPS offers a new route for adding protection to non-creative databases. To add such protection, not only could TRIPS be amended to include a provision that protects non-creative databases but also, since TRIPS incorporates the Berne Convention by reference, an amendment to the Berne Convention could accomplish the same.

In summary, although TRIPS provides the greatest international protection for databases through a multilateral treaty or trade agreement, assuming Poisindex is a non-creative database, even TRIPS fails to offer any protection. However, several regional agreements now extend protection to non-creative databases.

B. EU Database Directive

In 1996, the European Union adopted the Directive on the Legal Protection of Databases (EU Database Directive or Directive). This directive provides a double layer of protection. The first layer, covering creative databases, is grounded in copyright law, whereas the second layer, covering non-creative databases, is anchored in a sui generis right. The rationale for this bipartite protection was to make it easier to provide protection

116. See TRIPS art. 10(2).
117. Id.
118. See id. Part III.
119. See id.
for non-creative databases by avoiding any conflict with protecting certain groups of facts in copyright law.

First, Chapter II of the Directive provides copyright protection in a manner similar to that of the U.S. Copyright Act, the WIPO Treaty, and TRIPS. Article 3(1) requires originality in the sense that the database must be a collection of works or materials that by reason of their selection or their arrangement constitute an author’s own intellectual creation. However, this protection does not extend to the content of databases.

Second, Chapter III steps outside copyright law to create a sui generis right in certain compilations of data regardless of any creative organization. The important features of this protection include the following:

Protection of Databases Under the EU Database Directive: Present

| Source | The EU Database Directive is a regional agreement that protects non-creative databases.123 |
| Scope | Protection applies to (1) a collection of independent works, data, or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means,124 (2) for which the creator qualitatively or quantitatively invested substantially in obtaining, verifying, or presenting the contents.125 |
| Rights | The statute prohibits the extraction126 or re-utilization127 of the whole or of a substantial part of the contents of that database, evaluated qualitatively or quantitatively.128 |
| Exceptions | The rights are limited by numerous exceptions for member states, including (i) the extraction of nonelectronic databases for “private purposes”; (ii) the extraction for teaching or scientific research, but only if proper credit is given and the use has a noncommercial purpose; or (iii) the extraction or re-utilization is for the purpose of public security or an administrative or judicial procedure.129 These exceptions are narrower than the similar fair use exceptions under copyright law. For example, the sui generis right has no exceptions for criticism, news reporting, satire, or library use.130 However, the sui generis right is not subject to compulsory licensing arrangements, even in cases where the database compiler is the sole source of the database contents. |
| Duration | The property right in the database would continue for a fifteen-year period, but this period may restart upon any significant change to the database.131 |
| Treatment | The agreement requires reciprocity.132 |
| Enforcement | Enforcement occurs through each individual country’s laws, and can vary from country to country. |
| Remedies | Each member state is free to determine its remedies for the infringement of the sui generis right.133 |

122. See generally Caviedes, supra note 82.
124. Chapter I defines “database” as “a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means.” Database Directive, supra note 120, art. 1. “The requirement of independence is intended to exclude works such as a film, which literally is composed of a collection of still frames, but in which the still frames are not independent of each other.” Conley et al., supra note 23, at ¶ 81. Nonetheless, several commentators think the language is ambiguous. The “language of the Directive, however, leaves the scope of protection relatively uncertain at this time.” Schneider, supra note 121, at 556.
125. Database Directive, supra note 120, art. 7(1).
126. Id. art. 7(2)(a).
An “extraction” involves either the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form. Even the viewing of database contents on-screen constitutes an
Micromedex's Poisindex certainty falls within the scope of the EU Database Directive. However, since Micromedex is a U.S. database maker, it is doubtful that it could use the Directive to protect Poisindex in EU member countries.134

In summary, the current international framework provides virtually no international protection for U.S. non-creative databases. Neither the Berne Convention nor the WIPO Treaty apply to non-creative databases. Moreover, even if they did apply, neither treaty would enforce the rights of U.S. database makers. TRIPS boasts the authority through which U.S. database makers could enforce their rights, but it does not apply to non-creative databases. Finally, although the EU Database Directive applies directly to non-creative databases, U.S. database makers cannot use it to protect their investments in EU member countries because the United States does not offer comparable protection.

III. A Compromise Proposal for U.S. Database Protection

Congress should adopt new legislation to protect non-creative databases with an eye toward international protection. In this Part, the Note argues three propositions. First, although many commentators oppose any added protection for databases, not only is additional protection more innocuous than commentators threaten but also some form of protection is warranted to stimulate the innovation of useful databases. Second, both House Bill 354, which would give strong protection modeled on the EU's sui generis right, and House Bill 1858, which would give weak protection based on misappropriation law, harbor major shortcomings. Not only does House Bill 1858 fail to provide sufficient domestic protection for U.S. databases but also fails to procure needed international protection for U.S. database

---

127. Database Directive, supra note 120, art. 7(2)(b). "Re-utilization" means any form of making available to the public all or a substantial part of the contents of a database by the distribution of copies, by renting, or by other forms of transmission, including on-line." Schneider, supra note 121, at 559.

128. This right is much broader than it would at first appear, since the EU Database Directive expressly provides that the repeated and systematic extraction or re-utilization of insubstantial parts of a database may amount to the extraction or re-utilization of a substantial part. Database Directive, supra note 120, art. 7(5). One commentator points out that "[t]his is an important distinction, since the typical use of a database involves this very type of access." Conley et al., supra note 23, ¶ 83.

129. Database Directive, supra note 120, art. 9.

130. Conley et al., supra note 23, ¶ 84.

131. Database Directive, supra note 120, at art. 10(1)-(3). Article 10 establishes a term of protection for the sui generis right of fifteen years after the January of the year following the date when the database was first made available to the public or when any substantial change is made to the database. Id.; see also Sanks, supra note 22, at 999.

132. Database Directive, supra note 120, art. 11; see infra Part III.B.1.b.

133. Database Directive, supra note 120, art. 12.

134. For a detailed discussion explaining reasons U.S. database makers do not benefit from the EU Database Directive, see infra Part III.B.1.b.
makers. House Bill 354 is flawed because it would offer too much protection for database makers. Third, a compromise proposal can be designed to encourage the creation of innovative databases in the United States without impeding the access of data by the scientific community and the public in general, while, at the same time, ensuring that the United States does not fall behind other countries in the new global information-based economy.

A. An Argument for Additional Protection

Over the past several years, many commentators and powerful lobbyists\(^{135}\) have successfully urged Congress not to add protection for non-creative databases. However, the argument against additional protection is unfounded. None of the proposals give a monopoly over knowledge, alternative forms of protection are inadequate, and there is no constitutional bar to new protection.

1. The Argument for Protecting Non-creative Databases

Some form of U.S. legislation protecting non-creative databases is necessary: (a) to create an incentive for private compilers to make useful databases; (b) to ensure that someone is accountable for the accuracy of data in databases; and (c) to protect U.S. economic interests in a global information economy.

a. To Create an Incentive to Invest in New Databases

Without some form of protection, there is no incentive for individuals or companies such as Micromedex to take on the arduous and expensive task of producing useful databases that benefit the public.\(^{136}\) Commentators, courts, and database companies recognize this problem. One commentator writes that

[with such a thin layer of protection [provided by current U.S. law], the threat of piracy may discourage the development of commercially valuable databases . . . . Under this school of thought, the consumer is the ultimate victim of the Feist decision, because no one will be willing to compile the data that the consumer needs.\(^{137}\)]

\(^{135}\) Lobbyists against any added database protection include: (i) academic institutions who worry they would “be forced to pay for information that is currently free or inexpensive,” Kornblut, supra note 1, at A11; (ii) the scientific community, Sanks, supra note 22, at 993 n.19; (iii) Internet service providers who want to keep access to databases as free as possible, as part of their strategy to encourage people to do business online, Kornblut, supra note 1, at A11; (iv) technology firms—for example, on-line stockbrokers—who worry that if Congress goes for strong protection, the New York Stock Exchange would claim its stock quotes are part of a protected database and charge exorbitant licensing fees for each stock price, id.; and (v) libraries, including the Library of Congress, Sanks, supra note 22, at 993 n.19.

\(^{136}\) Producing databases is incredibly expensive. For example, Reed Elsevier states that its company has spent more than $650 million building Lexis-Nexis, a database of news, legal, and business information. Ron Eckstein, The Database Debate, LEGAL TIMES, Jan. 24, 2000, at 16, 17.

\(^{137}\) Wayman, supra note 84, at 434-35.

The majority's holding [that denied copyright to a database] establishes a rule of law that transforms the multi-billion dollar classified publishing industry from a business requiring the production of a useful directory based on multiple layers of creative decision-making, into a business requiring no more than a successful race to a data processing agency to copy another publisher's copyrighted work-product.139

Most importantly, companies recognize the risk of investing in data compilation. Although the Micromedex example is fictional, real cases of database piracy exist. For example, two high-tech firms provided consumers with a database containing home listings and other information at an Internet site.140 The site “allowed consumers to search for properties and real estate prices across the country, research potential new neighborhoods, find tips on selling their home, and even locate the fastest way to a home mortgage.”141 Although the services were initially successful, the success came at the expense of the National Association of Realtors (NAR):

The 700,000-plus-member organization [the NAR] says this Web site is one of many that are stealing information from the Multiple Listing Service, the national database of properties listed by Realtors from coast to coast. The NAR wants to quash what they see as data pirates siphoning business from those who worked long and hard to put the original database together.142

How many companies, individuals, and organizations have to lose their investments in databases before the entire industry recognizes the risk and halts further investment? Some evidence already shows that “the growth rates for both [databases and database producers] have slowed considerably in the years following the [Feist] decision, a signal of dampened investment in the industry.”143

Technological advances exacerbate the problem. Not only is the threat of piracy greater today than ever before, there are now more manifestations of piracy.144 The digitization of information poses special problems for protecting intellectual property.145 Today, pirating competitors can copy

---

138. 999 F.2d 1436 (11th Cir. 1993).
139. Id. at 1471 (Hatchett, J., dissenting).
142. Id.
144. See Conley et al., supra note 23, ¶ 10.
145. Id.

In the case of digital property, however, the temptation [to steal] is greater by many orders of magnitude than in the instance of, say, a patent on an industrial process or a copyrighted book. The ease with which digital property can be located, accessed, copied, modified, and distributed is utterly without precedent.
entire databases within minutes and distribute them effortlessly across the Internet.\textsuperscript{146}

Moreover, database makers are not only at risk to lose their investment to pirating competitors but also to "information samaritans"—people who extract data from a database without paying for it and make it available to the public for noneconomic reasons.\textsuperscript{147} This seemingly innocent infringement can cost manufacturers millions of dollars.\textsuperscript{148}

The point is simple. Databases serving legitimate public needs may never be realized without some economic incentive for their creation. Moreover, no economic incentive exists without some legal means to protect databases from piracy. Since the present law lacks sufficient protection,\textsuperscript{149} new legislation is needed. Following this reasoning, the current lack of database legislation ultimately hurts consumers\textsuperscript{150} and the progress of science.\textsuperscript{151}

b. To Enhance Accountability for Database Proprietors

An often overlooked fact is that database protection provides for the accountability of data suppliers. Without some form of ownership in data, data cannot be associated with specific distributors. As a result, the provider of information to a database cannot control its dissemination.\textsuperscript{152} Moreover, no mechanism exists to hold sloppy pirates accountable for providing false information. For example, a "parent browsing the Internet for poison remedies would have no legal recourse if the pirated version of the [Micromedex Poisindex], published without updates, revisions or accurate instructions, [or] gave outdated medical advice."\textsuperscript{153} This poses a real concern in an era where the Internet provides both true and false information indiscernibly coexisting side-by-side. The ultimate consequence is less accurate data in databases.

c. To Protect the U.S. Economy

Without additional protection for non-creative databases, the U.S. economy will suffer. The Information Age requires new legal developments.

\begin{itemize}
\item Id.
\item 146. Nelson, supra note 143, at 455. ("[A]ll electronic databases have two things in common: they are costly to produce, but they are easy to reproduce or copy.").
\item 147. Bastian, supra note 121, at 429.
\item 148. Tessensohn, supra note 57, at 482.
\item 149. See infra Part III.A.2.b.
\item 150. Wayman, supra note 84, at 435.
\item 151. Bastian, supra note 121, at 430 ("Consequently, as essential tools for improving productivity, advancing education and creating a more informed citizenry, a lower supply of databases can have a strong impact on technological progress.").
\item 152. Dennis R. Cronk, a realtor and President of the National Association of Realtors stated that "a pirating entity has no contractual or legal obligation to the real estate professional or consumer. There is no guaranteed protection of identity or addresses." Tighter Protection Against Piracy of Online Data: Top NAR Legislative Priority, PR Newswire, Jan. 19, 2000.
\item 153. Kornblut, supra note 1, at A11.
\end{itemize}
Information is the raw material for the new economic era which we entered upon some few short years ago. In the Agrarian Age, the law developed to facilitate the ownership and use of the most important asset of the time: land. In the Industrial Age, the law developed to facilitate the ownership and use of the chattel. So, in the Information Age, we should expect the law to develop to address the use of information, and in doing so to consider whether it can indeed be owned at all.¹⁵⁴

Economic analysis bears out the significance of the new global information-based economy: “Globally, expenditures in the information and communication technology markets exceed $1.8 trillion annually and comprise six percent (6%) of aggregate global Gross Domestic Product.”¹⁵⁵

Moreover, in this Information Age, nearly any form of data is potentially valuable. This includes business data (e.g., stock quotes, mutual fund reports, Securities and Exchange Commission filings); legal and government data (e.g., Federal, state and international court rulings, statutes, administrative regulations, proposed legislation, voter registration records, tax rolls); personal data (e.g., credit reports, buying preferences, student test scores, directories, health records, home listings, customer accounts, inventories, payrolls, the contents of art collections); scientific and technological data (e.g., the human genome, medical information); and entertainment data (e.g., sports statistics).¹⁵⁶ Various industries in the United States try to take advantage of the value in data, and although the sales in such industries have decreased since Feist, annual sales of database services and products still range from $4.5 billion to $200 billion in a given industry.¹⁵⁷ Databases play a significant role in the new global information-based economy.¹⁵⁸ Therefore, U.S. database protection is necessary to keep a dominant foothold in the information market.

2. Responding to the Opponents of Database Legislation

One commentator hypothesized that since “the information industry is growing explosively under the present dispensation,” no real need for legislation protecting non-creative databases exists.¹⁵⁹ “[A]ll the evidence suggests it ain’t broke, so don’t fix it.”¹⁶⁰ This myopic view of the database dilemma misses two major points. First, even if the U.S. database industry is surviving under the current lack of legal protection, it indisputably suf-

---

¹⁵⁴  Rees & Charlton, supra note 82, at 1.
¹⁵⁵  Conley et al., supra note 23, ¶ 1. “By 1997, such expenditures had escalated forty percent (40%) over 1992 levels.” Id. ¶ 1. Moreover, “revenues for the computer database industry have grown continuously at exponential rates, increasing from $9 billion to $19.2 billion between 1990 and 1994.” Wolken, supra note 6, at 1267-68.
¹⁵⁶  Nelson, supra note 143, at 453.
¹⁵⁷  The businesses profiting from data include the publishing industry and related services, newspapers, books and magazines, data processing and preparation, network services, business information suppliers, electronic delivery of business information, information retrieval services and commercial nonphysical research. Nelson, supra note 143, at 453 n.1.
¹⁵⁸  For figures on the international information industry, see supra Part II.
¹⁵⁹  Conley et al., supra note 23, ¶ 110.
¹⁶⁰  Id.
fers from a lack of protection. Fewer private investors and companies are willing to take on the risks associated with producing expensive, yet unprotected databases. Second, the present success of the U.S. database industry is not an indication of the industry’s health at the international level. Only a comparison between the United States and countries that protect their databases through legislation could alleviate worries at the international level, showing that the United States is not falling behind these countries. The suggestion that things are not so bad is shortsighted.

But opponents to U.S. database legislation also claim that (a) any legislative protection will create a monopoly over knowledge and will have long-term negative effects on science and public welfare, (b) any legislative protection is unnecessary because alternative forms of nonlegislative protection are available and adequate, and (c) stronger forms of legislative protection of non-creative databases are unconstitutional. Each argument is flawed.

a. Database Protection Will Create a Hazardous Monopoly over Knowledge

Many commentators claim that any additional database protection, especially strong protection modeled on the EU’s sui generis right, will indirectly hurt public access to information and impede the progress of science. The argument runs as follows. Creativity relies on a rich public domain of information because “[c]ulture, like science and technology, grows by accretion, [where] each new creator [builds] on the works of those who came before.” Increased protection in databases will allow a private monopoly over knowledge, and “[d]atabase producers would be able to restrict access to information . . . which would increase the cost of research . . . .” Consequently, higher prices for using data will keep information from scientists and the public. Although the argument for a sui generis right is logically valid, it is not sound for these reasons.

This argument against additional legislation contains three false assumptions. The above argument: (i) assumes that data, scientific or otherwise, will only be available through protected databases, which is either false or too speculative to warrant consideration; (ii) assumes incorrectly that a particular item of data must be tied only to a single database, which is false because all legislative proposals leave room for new database makers to create databases from scratch even though the final product is identical to a database already protected by legislation; and (iii) assumes numerous exceptions and additional legal remedies, such as antitrust laws, would not preclude the possibility of debilitating monopolies over data.

161. See generally Tessensohn, supra note 57.
162. Tessensohn, supra note 57, at 489.
If data are available through their original source, such as an individual researcher or a scientific article, then increased database protection will not block information from scientists and the public. Moreover, every indication suggests that data will continue to be available through their original source even if databases obtain added legislative protection.

Several commentators claim that "the database industry is heavily characterized by sole source data providers," where "sole source data' is defined as data that cannot be independently created, collected or obtained from any other source." This claim is unsupported. Database compilers must obtain information from original sources, and these original sources will not disappear just because a database incorporates data from them. Admittedly, there is the long-term possibility that if databases obtain legislative protection and become lucrative investments, database compilers may solicit data directly from researchers and original sources to preclude access to the data except through their database. In this case, the monopolistic effects of added protection could become real. However, this is presently not the case, and to suggest otherwise is speculative.

Furthermore, even in the worse case scenario, where a database provider is the sole source of a set of data, one must compare the cost of this situation with the value of knowing where to access the data. It is too simple to compare the case where data is tied to a monopoly and the case where data is free. Databases provide a more recognizable warehouse through which people may find data, and this convenience should be factored into the analysis. Data that is undiscoverable is useless to the public.

Second, if data are available through multiple databases, then presumably there exists competition, which will drive down the exorbitant prices anticipated by the opponents to legislation. Further, under every legislative proposal for protecting non-creative databases, any person is able to create a database from the original sources or from scratch without violating the law, even if the new database is identical to one legislatively protected. In other words, any database monopoly is intrinsically limited because competitors can go through the same procedure of compiling the


166. Not only is this speculative, but as Reichman himself admits, databases involve downstream uses of data—they involve completely digested data. Reichman & Uhlir, supra note 163, at 816-17. Thus, there will always be journal articles and other undisgested sources of information containing the same data that is in the public domain.

167. A surprisingly high number of commentators argue that the data in scientific experiments cannot be recreated except at a high cost. Reichman & Uhlir, supra note 163, at 807-08, 835. But this observation is irrelevant because the journal articles and experiments themselves remain available to other scientists.

168. E.g., H.R. 354, 106th Cong. § 2 (1999) (proposing 17 U.S.C. § 1403(c)) (stating "[n]othing in this chapter shall restrict any person from independently gathering information or using information obtained by means other than extracting it from a collection of information gathered, organized, or maintained by another person through the investment of substantial monetary or other resources").
rough data to create an identical database.\textsuperscript{169}

Third, added protection will not cause the demise of science because counter-balancing legal doctrines exist. Not only does every proposal for added protection include numerous exceptions to protect researchers and the public in certain contexts\textsuperscript{170} but also antitrust laws may address problems that arise.\textsuperscript{171}

In summary, no foundation exists for the argument that new legislative protection of non-creative databases will create a monopoly over knowledge and will have long-term negative effects on science and the public. Moreover, the manifesto echoing among anti-protection lobbyists that database legislation will allow the ownership of facts\textsuperscript{172} is folly. The only logical conclusion following from the opponent's argument is that data within a given database will be restricted, not that data themselves will be restricted. Ideally, added protection will encourage the investment and creation of novel databases, adding optional resources for researchers. Even if access through the new databases brings a price, it is better than no database at all. Opponents to database legislation are unrealistic to expect the best of both worlds.

b. Alternative Forms of Protection Already Exist to Protect Non-creative Databases

A second type of argument made by opponents to database legislation is that adequate alternative forms of protection are available.\textsuperscript{173} In particular, opponents argue that either (i) contractual protections, (ii) technological protections, or (iii) state misappropriation laws will suffice to protect database investments.\textsuperscript{174} Each suggestion is misguided.\textsuperscript{175}

\textsuperscript{169} Reichman & Uhlir, supra note 163, at 817-18. In response to this argument, one opponent to database legislation comments that "many databases cannot be recreated from scratch. Data that are time-sensitive, unique, very old, or prohibitively expensive fit this description. In research, this includes virtually all observational data sets of transient natural phenomena, as well as data from very costly or labor-intensive experiments." \textit{Id.} However, the fact that the process of compiling information is expensive does not mean that other ambitious entrepreneurs could not legally re-create any database from scratch. Critics of database protection want the best of both worlds; they want companies to invest millions of dollars to develop databases, and they also want to take advantage of this free labor.

\textsuperscript{170} See supra Part I.B.

\textsuperscript{171} Note that the provisions of the EU Database Directive "are without prejudice to the application Community or national competition rules." \textit{Sterling}, supra note 120, at 680. Moreover, under H.R. 354, "Nothing in this chapter shall limit in any way the constraints on the manner in which products and services may be provided to the public that are imposed by Federal and State antitrust laws, including those regarding single suppliers of products and services." H.R. 354 § 2 (proposing 17 U.S.C. § 1405(d)). \textit{But see} Tessensohn, supra note 57, at 480

\textsuperscript{172} Eckstein, supra note 136, at 16. Eckstein quotes Kevin Sheekey, Bloomberg Financial Markets' Washington representative, as saying "[i]f we allow people to own facts, at best we are putting a tollbooth onto the free flow of information and the Internet. At worst, we are building a dam." \textit{Id.}

\textsuperscript{173} E.g., Wayman, supra note 84, at 436.

\textsuperscript{174} Sanks, supra note 22, at 1010.
First, "after Feist, database owners strengthened their contractual protections by restricting the manner in which the database is used. For example, databases accessed through the Internet may contain restrictions for downloading and redistributing the database's contents." Moreover, licensing arrangements look even more promising in light of a recent court decision holding that licensing agreements are not pre-empted by the Copyright Act. Thus, "[a]lthough the copyright law itself will not protect the facts contained in a database, database providers can use shrink-wrap or click-wrap licenses to limit the access to and copying of the information in their databases."

However, there are numerous problems with relying on licensing agreements to protect databases. First, database owners have no reasonable means to monitor or police such agreements. Second, even though U.S. courts suggest that licensing agreements are not pre-empted by the Copyright Act, no guarantee exists that the typical shrink-wrap contract will be held enforceable by all courts. The ProCD Inc. v. Zeidenberg decision is not a guaranty that contractual limits on the use of mass-marketed databases will be uniformly enforceable. Third, "for many database proprietors, it will be impossible to obtain assent to restrictive contracts, or inconsistent with their business objectives. Fourth, remedies available for the breach of an agreement are inconsequential compared to copyright infringement remedies and remedies contemplated by legislative protection. And fifth, the costs associated with monitoring, draft-

Another option is to develop a more creative method of putting the database together to meet the Feist minimum level of creativity requirement. ... [But d]atabase owners perceive that creativity will add costs to development, while end users do not want the additional, potentially irrelevant information. Furthermore, there is no guarantee that this creative approach will actually yield copyright protection.

Id. (footnotes omitted).

175. Some commentators suggest that alternative forms of protection will be more successful for certain kinds of databases. Conley et al., supra note 23, ¶ 33.

176. Sanks, supra note 22, at 1008 (footnotes omitted).

177. ProCD Inc. v. Zeidenberg, 86 F.3d 1447, 1455 (7th Cir. 1996). "A copyright is a right against the world. Contracts, by contrast, severally affect only their parties; strangers may do as they please, so contracts do not create exclusive rights." Stone & Pernick, supra note 30, at 19.


179. Sanks, supra note 22, at 1008 ("Because the information does not physically reside with the database owner once copied, the database owner has no way to insure that the licensing agreement is honored.").

180. ProCD, 86 F.3d 1447.

181. Conley et al., supra note 23. "Notwithstanding the decision in Pro-CD, it appears that the jury is still out as to whether shrink-wrap licenses can be used to protect computer databases in light of the conflicting views of the various circuits." Tessensohn, supra note 57, at 457.

182. Conley et al., supra note 23, ¶ 70.

183. Tessensohn, supra note 57, at 458.

Principally, while specific enforcement of a contract is rarely available in action for breach, injunctive relief is standard in copyright infringement suits and operative throughout the country. Furthermore, a plaintiff in a breach of contract actions must prove damages, whereas copyright law provides statutory damages.
ing, and enforcing such agreements makes data more expensive for the public and researchers to access. In short, more contracts cannot solve the database dilemma.

Second, opponents to added protection argue that technological protections sufficiently fortify databases against pirates, but such protections are insufficient. First, “if database owners must resort to a technological solution, the cost of developing databases will increase.” Second, even if encryption technology was sufficiently advanced, which it is not, it “cannot protect databases once the contents are printed.” In summary, most commentators agree that “[d]espite their sophistication, self-help technological approaches like these have not eliminated the demand for legal protections. In some contexts, they are technologically infeasible. In other situations, they would impose a burden on existing and potential users that would be unacceptable from a business perspective.”

Third, opponents to added database protection sometimes suggest that current state misappropriation laws provide sufficient civil remedies against pirates. The doctrine of misappropriation arises out of International News Service v. Associated Press, where the Supreme Court held that a news agency could not reprint news gathered by another agency. Several commentators suggest that this doctrine could be used to protect databases. However, at present, there is no reason to think courts are willing to expand the doctrine to cover databases. First, the doctrine applies narrowly to “hot news”—time-sensitive information, not scientific research archives. Second, the test for applying this judicial doctrine is simply too uncertain for database makers to rely on it. Third, “[u]nfair

and the possibility of an award of costs and attorney’s fees to the prevailing party.

Id.

184. Technological protections include encryption technology. For a detailed discussion of current technological devices for protecting the content of databases, see Conley et al., supra note 23, ¶¶ 26-33.

185. Sanks, supra note 22, at 1009.

186. Id.


188. Fair Use of Databases, N.Y. Times, Nov. 15, 1999, at A26 (“Databases already enjoy some protections. Many states have misappropriation laws that can be used to bar database piracy, while federal copyright law affords protection to the expression, selection and arrangement of facts within a database, but not, it must be stressed, to the facts themselves.”).

189. 248 U.S. 215, 239 (1918).


Not surprisingly, the most recent definition of misappropriation, set forth by the Second Circuit in National Basketball Association v. Motorola, Inc., interprets INS v. AP narrowly. The Second Circuit reasoned that INS v. AP should be held to its facts owing to a lack of an analytic underpinning for the doctrine of misappropriation.

Id. “Unfortunately, most databases are usually archival in nature and not ‘hot news.’” Tessensohn, supra note 57, at 462.

192. Tessensohn, supra note 57, at 463.
competition and misappropriation remedies are likely to face preemption" by the Copyright Act.\textsuperscript{193} In short, the unfounded suggestion that current misappropriation laws can protect databases is wishful thinking.

Therefore, although current contractual, technological, and state misappropriation protections provide hints of security to database proprietors, they offer only the faintest hope of protecting huge investments in data compilation and maintenance.

c. Database Protection Is Unconstitutional

A third argument commonly offered against stronger forms of legislative protection for databases\textsuperscript{194} is that such legislation is unconstitutional. The Supreme Court has already held in \textit{Feist} that labor alone cannot supplant the creativity requirement.\textsuperscript{195} This argument is hardly worth mentioning because legislators claim that any new database protection would fall under the Commerce Clause\textsuperscript{196} if not the Intellectual Property Clause.\textsuperscript{197} In fact, even if new legislation is manifested as an amendment to the Copyright Act, the fact that the Commerce Clause, not the Intellectual Property Clause, license it is insubstantial.\textsuperscript{198} Provisions in the Copyright Act need not rest on the Intellectual Property Clause.

Beyond the obvious need for some form of added protection for non-creative databases, opponents to added protection offer no cogent reason not to provide such protections. Therefore, the issue turns to which form of legislation is most suitable for resolving the database dilemma at the national and international levels.

\textsuperscript{193} Conley et al., \textit{supra} note 23, ¶ 70.

To survive [preemption], such remedies cannot be merely 'copyright like' causes of action, but must include extra elements such as time-sensitivity of the data or free-riding. However, what many database proprietors really want is precisely what the preemption doctrine forbids: a simple prohibition against copying compilations that are insufficiently original for copyright purposes.

\textit{Id.}

\textsuperscript{194} The argument does not apply to weaker forms of database protection modeled on misappropriation law such as H.R. 1858, 106th Cong. (1999).


\textsuperscript{196} U.S. CONsT. art. I, § 8, cl. 3.

\textsuperscript{197} \textit{Id.} art. I, § 8, cl. 8; see Nelson, \textit{supra} note 143, at 474-75.

\textsuperscript{198} A few commentators suggest that database protection would violate the First Amendment.

For some forty years, the late Professor Melville Nimmer, a leading authority on both copyright and First Amendment law, taught that copyright protection would violate First Amendment guarantees of free speech were it not for the judicial exclusion of ideas and facts from the reach of the exclusive property rights granted to authors and artists.

Reichman & Uhlir, \textit{supra} note 163, at 833.
B. The Problem with the Current Legislative Proposals

Both of the recent legislative proposals for added protection, House Bill 1858 and House Bill 354, contain major shortcomings. While House Bill 1858 is too weak, House Bill 354 is unnecessarily strong.

1. The Problem with House Bill 1858

At present, the misappropriation approach to database legislation receives much support as a reasonable alternative to the strong forms of protection characteristic of early U.S. legislative proposals and House Bill 354. Nonetheless, House Bill 1858 contains serious flaws. House Bill 1858 is problematic because (a) it offers little or no real protection for database compilers in the United States, and (b) it neglects to create important international protections for U.S. database makers by failing to provide protection comparable to that outlined in the EU Database Directive.

a. Domestic Shortcomings

House Bill 1858 provides almost no real U.S. protection for database makers. Not only is the scope of protection narrow to the point of making it trivial but also the bill leaves no mechanism through which a victim of piracy may enforce her rights.

House Bill 1858 fails to proscribe any piracy except the most unimaginative, wholesale copying of databases. Assuming that the database was compiled through the "investment of substantial monetary or other resources," the proposal makes it:

unlawful for any person, by any means or instrumentality of interstate or foreign commerce or communications, to sell or distribute to the public a database that (1) is a duplicate of another database that was collected and organized by another person; and (2) is sold or distributed in commerce in competition with that other database.

According to the proposal, "a database is 'a duplicate' of any other database if the database is substantially the same as such other database, and was made by extracting information from such other database." The language fails to cover numerous forms of piracy-like conduct.

First, a pirate could extract substantial amounts of information from a database without violating the statute, so long as she does not ultimately create a new database substantially similar to the original database. For

---

199. Bastian, supra note 121, (recommending protection based on misappropriation); Tessinohn, supra note 57; Fair Use of Databases, supra note 188, at A26; MacMillan, supra note 164.
200. "H.R. 1858 will simply give broad leeway to other industries to pirate and sell the material that our companies have worked so hard to organize and bring to the marketplace." House Adjourns Without Taking Action On Database Protection Legislation, REP. ON DIRECTORY PUBLISHING, Dec. 1999, at Vol.9, No.12 LEXIS, News Library, SIMBA File (quoting Dan Duncan, vice president of government relations for the Washington-based Software Information and Industry Association (SIIA)).
202. Id. § 102.
203. Id. § 101 (emphasis added).
example, a pet store could simply download all of Micromedex's poison antidotes dealing with snakes and poisonous animals and incorporate it into a different pet-related database without violating House Bill 1858; the final product is not substantially the same as Poisindex.

It should be noted that House Bill 1858 defines "database" such that there may be databases within databases. The bill states that "a discrete section of a database that contains multiple discrete items of information may also be treated as a database."204 The significance of this is that Micromedex may be able to protect discrete subsets of data, such as antidotes to animal poisons within their database. However, the bill's language does not designate which subsets of data may constitute independently protectable databases. Moreover, the bill's language does not prescribe how a database proprietor may ensure the protection of these discrete subsets of data in a larger database.

Second, a large health services corporation could extract Micromedex's entire database and escape infringement by transforming it into a database that is not "substantially" the same as the original by simply adding data. Micromedex would lose its entire investment without remedies under House Bill 1858.

Third, House Bill 1858 leaves the victim of piracy helpless to pursue an action against the suspected pirate. Under the bill, the Federal Trade Commission has the exclusive authority to bring an action,205 which leaves the decision of whether to bring an action to its discretion, with the possibility for indefinite delays in the recovery against even an obvious case of misappropriation.206 The proposal states that "[t]he Federal Trade Commission shall have jurisdiction, under Section 5 of the Federal Trade Commission Act . . . to prevent violations of section 102 of this title."207 In short, a database provider cannot adequately protect its investment under the bill.

b. International Shortcomings

House Bill 1858 provides no additional international protection for U.S. database makers. This is apparent from two facts: (i) House Bill 1858 would fail to secure reciprocal sui generis protection from EU member countries, and (ii) current international treaties and trade agreements fail to offer any international protection for U.S. databases.

Citizens and companies in the United States cannot benefit from the sui generis rights of EU member countries because the EU Database Directive indirectly calls for reciprocity.208 Generally, "the Directive limits the

204. Id. (defining "database").
205. Id. §§ 107, 201.
208. This reciprocity requirement motivated U.S. legislators to consider H.R. 3531, which was modeled closely on the sui generis right of the EU Database Directive. Sanks, supra note 22, at 994.
sui generis right to database makers or rightholders who are nationals of a Member State or who have their residence in the territory of the Community. However, according to the Directive, a foreign entity may use the EU sui generis right to protect its database in an EU member country if one of two conditions are met. First, if the entity is a corporation or partnership, the entity must have been formed in "accordance with the law of a Member State and have their registered office, central administration or principal place of business within the Community." This provision fails to help U.S. database makers. Second, the sui generis right applies to databases "produced by legal persons not established in a Member State, within the meaning of the Treaty, only if such third countries offer comparable protection to databases by nationals of a Member State or persons who have their habitual residence in the territory of the Community." In other words, individuals and entities "outside the EU . . . may not claim the sui generis right unless they reside or were incorporated or formed in a jurisdiction which provides comparable protection for EU databases."

Unfortunately, the term "comparable protection" is left open to interpretation by the EU Council. In determining whether U.S. database legislation is comparable, the "EU Council may adopt a reciprocal arrangement upon a proposal from the Commission." Clearly, House Bill 1858 would not provide "comparable protection" and would thus fail to satisfy the European Union's reciprocity requirement. In short, House Bill 1858 fails to provide any added protection in EU member countries.

Second, the national treatment provisions in the Berne Convention, the WIPO Treaty, and TRIPS fail to protect U.S. database makers. Since both the Berne Convention and the WIPO Treaty require national treatment, an EU member could not deny Micromedex copyrights provided to its own citizens without violating the Berne Convention and the WIPO Treaty. However, if an EU member country denied Micromedex its own national sui generis right in databases, this apparently would not violate the Berne Convention or the WIPO Treaty, since both refer only to copy-
rights. In other words, the sui generis right falls outside the scope of these treaties. But even if the sui generis right fell within the scope of either treaty, a violation of the treaty poses no consequences to an EU member country that denied Micromedex its sui generis right. In other words, the treaties are doubly ineffectual for securing international database protection.

At first blush, although TRIPS offers a different hope because it provides enforcement measures, national treatment, and most favored nation treatment, it nonetheless fails to supply international protection. Like the Berne Convention and the WIPO Treaty, TRIPS affords no protection under its national treatment provision. Notably, the TRIPS provision dealing directly with databases seems to encompass a copyright in non-creative databases, which would allow Micromedex to enforce the copyright in an EU member state. Unfortunately, the national treatment provision in TRIPS only applies to copyright laws, not to a sui generis law protecting databases. Again, Article 10(2) states that even though protection does not extend to the actual data contained in the compilations, it is "without prejudice to any copyright subsisting in the data or material" that a member state may choose to grant. The EU sui generis right goes outside copyright to protect databases.

The most favored nation treatment provision in TRIPS fails to apply to the EU sui generis right for similar reasons. TRIPS provides that "any advantage, favour, privilege or immunity granted by a Member to the nationals of any other country shall be accorded immediately and unconditionally to the nationals of all other Members." Although "[t]his means that in the future, the grant of intellectual property rights cannot be based on a requirement of reciprocity," again, the sui generis right in the EU Database Directive falls outside the scope of TRIPS, thus leaving TRIPS incapable of covering database protection at the international level.

Some commentators claim that no additional international protection is needed because U.S. databases are currently thriving without international protection. However, this claim pales in the light of several observations.

First, U.S. companies are currently losing millions of dollars to foreign pirates of copyrightable works. Although no data on foreign losses attributed solely to database thievery exists, "[t]he International Intellectual Property Alliance estimated that in 1998 losses [due to piracy] were about $5 billion for businesses." Moreover, "for business applications, [a loss]
over $3 billion . . . creates distortions in the market and creates parallel black market systems which in the end will hurt, not just the [United States], but the global economy as a whole."223 In short, international piracy is a real concern.

Second, the United States has the most to lose in the new global information-based economy because of a lack of international intellectual property protection since it is the single largest producer of database services and products. Typically, the strongest resistance to increased international intellectual property protection comes from countries holding the fewest intellectual property products. The rationale for such countries is that they have nothing to lose from little protection—they have no product to steal—and everything to gain from little protection—they can steal the products of other countries with impunity. In the case of database protection, the United States has the most to lose and nothing to gain from a lack of international protection.

A third and related reason for international protection is that U.S. database makers are currently at risk of falling behind foreign database makers. Most commentators agree that without legislation comparable to the EU sui generis right, U.S. database makers will suffer a competitive disadvantage.224 Since competing countries in the European Union and elsewhere protect their companies' database investments, the United States is disadvantaged in two ways. Not only can European companies succeed where U.S. companies fail but also European companies may filch the investments of U.S. database providers without facing penalties.

Fourth, there is little doubt that the European Union has set the agenda for international database protection by requiring a reciprocity provision in its Database Directive.225 In other words, although the United States may resist increased protection for non-creative databases, they are certainly in the minority.226 Long-term resistance would come at the price of economic isolation.

In summary, House Bill 1858 is unacceptable. Not only does it fail to provide sufficient protection for U.S. database makers within the United States but also it cannot extend database protection internationally. Moreover, some form of international protection is crucial to continued prosperity for U.S. database makers.

223. Id.
224. Schneider, supra note 121, at 561 (citing Anna Kraske & Thomas A. Unger, Protecting Works on CD-ROM and Internet, 2 INTELL. PROP. STRATEGIST 8, 9 (1996)); Tessensohn, supra note 57, at 466 ("The bottom line is that many U.S. computer databases that are worth billions of dollars may be denied protection under the EU Directive, placing U.S. producers at a competitive disadvantage throughout the huge European market . . .").
225. Schneider, supra note 121, at 561.
2. The Problem with H.R. 354

Unlike House Bill 1858, most commentators criticize House Bill 354. Although a number of the criticisms are unfounded, House Bill 354 is defective because it offers too much protection to U.S. database makers in the United States. This overprotection includes: (a) the possible protection of small databases, which could revive a form of the sole-source problem; (b) possibly preventing investors from creating new databases by combining parts of others that are protected; (c) the possibility that database makers could force consumers to contract around exceptions to House Bill 354; (d) an effectively perpetual term of protection; (e) potentially overcautious infringers likely avoiding data use that is otherwise legally accessible because the exceptions to protection place the burden of proof on the alleged infringer; and (f) the vagueness and imprecision of terms and exceptions that may cause a chilling effect in the scientific and academic communities.

a. The Sole-Source Problem

Since House Bill 354 provides no guarantee that small databases are precluded from protection, original sources of data may claim protection under the bill, thus presenting a version of the sole-source problem. The bill provides protection to "a collection of information gathered, organized, or maintained by another person through the investment of substantial monetary or other resources...." The bill defines "collection of information" to mean "information that has been collected and has been organized for the purpose of bringing discrete items of information together in one place or through one source so that users may access them." Under a reasonable interpretation of these provisions, it is conceivable that a researcher combining fifty items of scientific data in a single research article could claim protection under House Bill 354. The researcher may argue that her single article brings discrete items of information together in one place. Moreover, the researcher may argue that because the data was acquired through elaborate, time-consuming and expensive experiments, the data was gathered "through the investment of substantial monetary or other resources." Unless an additional provision in the bill confines protection to databases compiled from original sources, small original sources may be monopolized.

227. For example, one commentator writes: "[T]his paper will expose the dangers of the second initiative, The Collections of Information Act, which is currently pending in the House of Representatives." Tessensohn, supra note 57, at 440-41.
228. For example, many commentators make sweeping claims against creating a property right over knowledge. See supra Part III.A.2.
229. See supra Part III.A.2.
232. Id. (proposing 17 U.S.C. § 1402). Hopefully, a court interpreting the statute would restrict the "sweat of the brow" test to acquiring data from original sources rather than from discovery itself. But there is no guarantee that courts will interpret House Bill 354 in this way.
b. The Re-compilation Problem

Potentially, House Bill 354 precludes the re-compilation of data from protected databases, "balkanizing" data compilations. Discussing a predecessor of House Bill 354, one commentator claimed that under strong forms of legislative protection,

[|later scientists and engineers could not combine data legitimately accessed from one commercial database with data extracted from other databases to make a complex new database for addressing hard problems without obtaining additional licenses and permissions. This remains, perhaps, the single most critical problem for scientific and technical research. No one could combine "substantial" amounts of data or information into a more efficient follow-on product without a license; the licensor would labor under no duty to grant such a license; and the sole-source provider would not want any competition from follow-on products.]

The idea is that once databases obtain protection, inventive and useful combinations will become expensive and difficult. For example, if an epidemiologist wants to combine the data of fifty private, protected databases in a new database to study the cause of a deadly disease, it would be prohibitively expensive to compile the new database from scratch. Moreover, the protected databases may hold out or refuse to license the use of the information to maintain their monopolistic position in the market. In short, some commentators surmise that under protection like House Bill 354, data will become balkanized, impeding scientific advancement.

c. Freedom of Contract Issues

The exceptions listed in House Bill 354—designed to protect educational, scientific, and other uses—are less helpful than they initially appear because database makers may be able to force consumers to contract around these exceptions.

[The ability of data providers to override by contract even the limited exceptions that the new law may grant to public-interest users, including scientists, engineers, and educators, is of great concern. Without a concomitant duty to deal fairly and reasonably with public-interest users, these combined powers could lead to high prices for data and to the imposition of harsh and oppressive terms concerning both access and subsequent uses of data that would especially disadvantage academic researchers.]

The law of contracts is not affected by the database protection provided under House Bill 354. Moreover, although House Bill 354 preempts state law providing equivalent rights, the bill states that "the law of contract shall not be deemed to provide equivalent rights . . . ." Because parties can override exceptions through contract, the exceptions

---

233. Reichman & Uhlir, supra note 163.
234. Id. at 808-09.
235. Id. at 814-15. The author notes that although this comment is directed at a predecessor of House Bill 354, it applies equally to House Bill 354.
236. See H.R. 354 § 2 (proposing 17 U.S.C. § 1405(a)).
237. Id. (proposing 17 U.S.C. § 1405(b)).
are less helpful than they first appear.²³⁸

d. The Perpetual Duration Problem

Although House Bill 354 would create a fifteen-year term for database protection, the protection of data could effectively become perpetual. House Bill 354 states that

[n]o criminal or civil action shall be maintained under this chapter for the extraction or use of all or a substantial part of a collection of information that occurs more than fifteen years after the portion of the collection that is extracted or used was first offered for sale or otherwise in commerce, following the investment of resources that qualified that portion of the collection for protection under this chapter.²³⁹

Suppose Micromedex publishes and sells its original version of Poisindex in 2000. In 2005, Micromedex collects additional poison antidotes, adds them to Poisindex, and sells the new version. The original collection of poison antidotes would be protected until 2015, whereas the added portion would carry protection until 2020. The problem is that even if a user of Poisindex in 2017 knows that most of Poisindex has entered the public domain, she probably will not be able to access and use the public domain portion for fear of violating the statute unless she can discern the protected portions from the unprotected portions. In short, Micromedex could effectively obtain perpetual protection of its entire database by mixing protected and unprotected data.²⁴⁰

e. Issues Related to the Burden of Proof

Since House Bill 354’s “permitted acts” exception is framed in terms of a defense, rather than an element of infringement, scientists and other users who are intended to be exempt from liability may nonetheless avoid using data in protected databases to avoid the possibility of costly litigation.²⁴¹ House Bill 354 provides a blanket exception to database protection: “No person shall be restricted from extracting or using information for non-profit educational, scientific, or research purposes in a manner that does not harm directly the actual market for the product or service . . . .”²⁴² In addition, House Bill 354 provides an exception for certain limited uses or extractions: “In general . . . an individual act of use or extraction of information done for the purpose of illustration, explanation, example, comment, criticism, teaching, research, or analysis, in an amount appropriate and customary for that purpose, is not a violation of this chapter, if it is

²³⁸. “Nothing in this chapter shall restrict the rights of parties freely to enter into licenses or any other contracts with respect to the use of collections of information.” Id. (proposing 17 U.S.C. § 1405(e)).
²³⁹. Id. (proposing 17 U.S.C. § 1408(e)) (emphasis added).
²⁴⁰. Reichman & Uhlir, supra note 163, at 809-10.
²⁴¹. Id. at 805.
²⁴². H.R. 354 § 2 (proposing 17 U.S.C. § 1403(a)(1)).
reasonable under the circumstances.”

Even if a scientist is clearly exempt from liability under this provision, House Bill 354 places the burden of proof on the scientist to demonstrate her exempt status, rather than the database maker to show her nonexempt status. As a consequence, clearly exempt users may completely avoid use to avoid the mere possibility of litigation. The result is that scientists would avoid using data even though such use is legally permissible.

f. The Problem of Indefiniteness

Any vague language or test within a statute creates uncertainty, where, depending on the statute's interpretation, a person may or may not be held liable. Consequently, good conservative legal advice would recommend staying out of clear violative conduct and conduct lying within the bounds of uncertainty, even if such conduct is legal. Uncertainty effectively augments the domain of conduct proscribed by the statute.

House Bill 354 contains several vague terms and at least one indefinite test. First, the expression “substantial part” in Section 1402 is vague. The statute reads: “Any person who extracts, or uses in commerce, all or a substantial part, measured either quantitatively or qualitatively, of a collection of information” may be liable provided certain additional conditions are met.

In describing certain permitted acts, the bill states clearly that the “extraction or use of an individual item of information” does not constitute a violation of the bill. Moreover, the statute specifically proscribes “the repeated or systematic extraction or use of individual items or insubstantial parts of a collection of information so as to circumvent the prohibition” under the general provisions of the statute.

Unfortunately, the elaboration of “substantial part” ends there. Thus, Section 1402 is too ill-defined to provide a real limitation on database protection.

Criticizing analogous language in the EU Database Directive, one commentator writes that “the risks of invoking even this exception are high, because a would-be user has no way of knowing in advance whether a court will later find that the amount used was in fact qualitatively or quantitatively insubstantial.” In particular, because the notion of a “collection of information” is not well-defined, a clever database owner may argue that the relevant “collection of information” consists of all the items of information taken by the alleged pirate, whatever those turn out to be. For example, in the case where a pet store copies all of Micromedex’s

---

243. Id. (proposing 17 U.S.C. § 1403(a)(2)(A)). The bill then goes on to describe four factors a court is to consider in determining whether an act of use or extraction is reasonable. Id. (proposing 17 U.S.C. § 1403(2)(a)(i)-(iv), (2)(b)).

244. Although this uncertainty is often eradicated by a rich common law interpreting the statute, such uncertainty is always significant at the enactment of a statute.

245. Id. (proposing 17 U.S.C § 1402) (emphasis added).

246. Id. (proposing 17 U.S.C. § 1403(b)).

247. Id.

248. Reichman & Uhlir, supra note 163, at 804.

animal poison antidotes, the pet store is liable for extracting the whole of Micromedex's database if Micromedex can persuasively argue that the relevant "collection of information" is animal poison antidotes and not all poison antidotes. The notion of a "substantial part" is relative to the working conception of a "collection of information." As long as a database owner can potentially define the database in an ad hoc fashion, scientists may shy away from even legal uses of the contents of protected databases.

Second, the expression "harm to the actual or potential market" in Section 1402 is vague. The statute states that a use or extraction of data "so as to cause harm to the actual or potential market of that other person" may be liable, provided that certain additional conditions are met. Unfortunately, House Bill 354 does not define the type or amount of harm needed to violate the statute. Since a single unpaid extraction from a database constitutes a loss of payment to the database owner, every use or extraction potentially harms the actual market. A liberal interpretation shows that almost any use or extraction of data from a protected database could meet the standard. The uncertainty surrounding the language could stifle otherwise legal uses and extractions.

Third, the expression "gathered, organized, or maintained by another person through the investment of substantial monetary or other resources" in Section 1402 is vague. The statute only protects a "collection of information gathered, organized, or maintained by another person through the investment of substantial monetary or other resources..." Again, this expression is not defined or further explained. Until a database proprietor successfully protects his database through litigation, there is no clear way to know whether the statute will protect the database.

Fourth, the blanket exception for individual acts of use or extraction for "Educational, Scientific, Research, and Additional Reasonable Uses" is too indefinite to ease anxiety among otherwise lawful users of protected databases. The exception states that

[i]n determining whether such an act is reasonable under the circumstances, the following factors shall be considered:

(i) The extent to which the use or extraction is commercial or nonprofit.

251. Note that the exception in Section 1403(a) for educational, scientific, research, and other reasonable uses states that an extraction or use in this context is violative only if the extraction or use harms "directly the actual market for the product..." Id. (proposing 17 U.S.C. § 1403(a)(1)). This is helpful in contexts such as nonprofit educational or scientific uses, but it does not help in other instances.
252. House Bill 354 defines "potential market." Id. Nevertheless, "would-be competitors would never know in advance when the use or extraction of protected data may turn out to cause harm to some unknown potential market." Reichman & Uhli, supra note 163, at 837.
254. For purposes of this paragraph, the term 'individual act' means an act that is not part of a pattern, system, or repeated practice by the same party, related parties, or parties acting in concert with respect to the same collection of information or a series of related collections of information.
Id. (proposing 17 U.S.C. § 1403(a)(2)(B)).
(ii) The good faith of the person making the use or extraction.

(iii) The extent to which and the manner in which the portion used or extracted is incorporated into an independent work or collection, and the degree of difference between the collection from which the use or extraction is made and the independent work or collection.

(iv) Whether the collection from which the use or extraction is made is primarily developed for or marketed to persons engaged in the same field or business as the person making the use or extraction.\(^\text{255}\)

This exception is modeled on the Copyright Act's fair use provision, which loosely codifies the common law exception for certain educational and research uses of copyrighted works. Although this exception potentially provides a broad shield to nonprofit researchers and users in educational contexts, the indefiniteness of the test is likely to cause otherwise legal users to avoid legal uses. The threat of litigation effectively chills use.\(^\text{256}\)

In summary, House Bill 354 is flawed because it unnecessarily offers too much protection for U.S. databases in the United States. It leaves potential sole-source problems; re-compilation problems; the problem of contracting around important exceptions; durational problems; burden of proof problems; and problems associated with uncertainty, including vague terms and indefinite tests. Better tailored legislation could address these problems.

C. A Compromise: Resolving the Database Dilemma with an Eye Toward International Projection

U.S. database legislation must strike a balance between House Bill 1858 and House Bill 354.\(^\text{257}\) The goal in the context of database protection is to provide enough national and international protection to encourage the compilation of large commercial databases while leaving room for reasona-

\(^{255}\) H.R. 354 § 2 (proposing 17 U.S.C. § 1403(a)(2)(A)(i-iv)). This exception also states that:

\[\text{In no case shall a use or extraction be permitted under this paragraph if the used or extracted portion is offered or intended to be offered for sale or otherwise in commerce and is likely to serve as a market substitute for all or part of the collection from which the use or extraction is made.}\]

\[\text{id.}\]

\(^{256}\) Note that the chilling effect may be mitigated by a provision in the civil remedies portion of the statute. Section 1406(e) states that

\[\text{[t]he court shall reduce or remit entirely monetary relief under subsection (d) in any case in which a defendant believed and had reasonable grounds for believing that his or her conduct was permissible under this chapter, if the defendant was an employee or agent of a nonprofit educational, scientific, or research institution, library, or archives acting within the scope of his or her employment.}\]

\[\text{id. (proposing 17 U.S.C. § 1406(a)). This gives the court broad discretion to limit damages in such cases.}\]

\(^{257}\) Indeed, as one commentator has written, "[t]he guiding principle has always been balance: dole out only as much exclusivity as is necessary 'to promote the progress of science and the useful arts,' but otherwise, protect the public domain." Conley et al., supra note 23, ¶ 8. This commentator went on to observe that "[a]rriving at a perfect utilitarian balance between incentives for knowledge producers and access for knowledge consumers has always been the Holy Grail of intellectual property law." Id. ¶ 93.
bly priced data access for researchers. Congress can only meet this goal by providing less protection than House Bill 354, the sui generis approach, and more protection than House Bill 1858, the misappropriation approach.

1. **The Compromise**

Each of the problems associated with overprotection in House Bill 354 can be mitigated by changing provisions in the bill. The provisional changes include: (a) creating an element for infringement that targets proprietors attempting to monopolize a sole source of data; (b) creating a compulsory license in certain circumstances; (c) creating a limited preemption of contracts that would otherwise surrender a user’s rights under the statutory exceptions to database protection; (d) raising the standard for obtaining a new term of protection; (e) switching the burden of proof to the plaintiff to show that the defendant does not have an exempt status; and (f) creating more definite rules through precise terms in the proposed statute and creating specific guidelines exempting certain educational, scientific, and research uses.

a. **Solving the Sole-Source Problem**

To solve the sole-source problem, a provision should be added to House Bill 354 requiring, as an additional element for proving a violation of the statute, that the database-proprietor plaintiff demonstrate that the database was not the sole source of the information extracted or used. Therefore, to prove a prima facie case of a violation of the statute, the plaintiff would have to show that each item in the database allegedly extracted or used was publicly accessible through alternative means.\(^\text{258}\)

A second provision should state that if the alleged pirate used or extracted more than seventy-five percent of the items in the database at the time of the alleged misappropriation, then the plaintiff would have to prove either that at least fifty percent of the items in the database were publicly accessible\(^\text{259}\) or that each item misappropriated was publicly accessible, whichever is less. The information must be publicly accessible because, otherwise, a database proprietor could solicit the original source, incorporate the data into her database, and therefore maintain a monopoly. Congress should define “publicly accessible” data as data the public can obtain through reasonable diligence from a printed or electronic source that is

\(^{258}\) As currently written, the other elements for a prima facie case of violation for H.R. 354 include: (i) that the plaintiff was the owner or rightsholder of the database; (ii) that the defendant extracted or used in commerce a substantial part of the database; (iii) that the plaintiff expended sufficient money or resources to gather, organize, or maintain the database; (iv) that the use or extraction caused harm to the plaintiff’s actual or potential market; and (v) that the plaintiff offered or intended to offer in commerce the database containing the information taken. H.R. 354 § 2.

\(^{259}\) Under this second provision, the plaintiff could show that any fifty percent of the items were publicly accessible, not just those that were misappropriated. The percentages are somewhat arbitrary and should be refined to reflect any empirical information on the subject.
open and free to the public or obtainable, as individual items of data, through inexpensive research.

For example, federal appellate opinions and scientific articles and treatises are publicly accessible through public and private libraries. Although some directories, such as the White Pages, are not free, the public can obtain the individual items of data through inexpensive research.\textsuperscript{260} It follows that if the public cannot obtain the data in the database from alternative sources and cannot reproduce individual data items at a low cost, the database would not gain protection.\textsuperscript{261}

The plaintiff's burden should not consist of showing that every item in the database was publicly accessible. The purpose of the provision is to ensure that the data in a protected database are available through alternative means. It is wasteful and unrealistic to expect the database proprietor to prove that every item in a database containing 10,000 items was available through alternative means, especially if most of that data are not relevant to the data that was allegedly misappropriated. Moreover, if the alleged pirate extracts the entire database or a substantial amount of it, the plaintiff's failure to prove that a single item in the database was not publicly accessible would not defeat her claim. Alternatively, the plaintiff could prove that at least fifty percent of the database was publicly accessible at the time of copying.\textsuperscript{262}

Requiring the database proprietor to prove that each item of the database that was wrongfully extracted was available through alternative means effectively ensures that every item in the database is available through alternative means. Under this provision, a database user knows either that the database is the sole source of the data and therefore not protected, or she knows there is some other publicly accessible source for the information. If the database user discovers data she would like to extract, she can always look for the same data from alternative sources. If she cannot find the data from a publicly accessible source after using reasonable diligence, she can assume that the database is the sole source of the information and not protected. The user must make a calculated decision, but if suit is brought, the burden of proof in litigation rests on the database proprietor to prove that the data were actually available from

\textsuperscript{260} The author draws a line between obtaining individual items of data and whole collections of data. If obtaining an individual item of data is expensive and is not obtainable through printed or electronic means for free, then it would not be publicly accessible. Thus, a scientist conducting a costly experiment and producing only ten items of data could not protect the items under the statute. Moreover, the law would not protect any database that is the sole source containing the items.

\textsuperscript{261} Nothing would prevent a database proprietor from buying a collection of data from a sole-source provider and then placing each item of data in such a diffuse fashion that a person could not obtain the data from publicly accessible sources in any convenient way. The database proprietor obtains the data as a packet—for a cost—but then creates an inconvenience for users to obtain the data elsewhere by spreading it out over public domain sources. Hopefully, the process of spreading out data over publicly accessible sources is too costly for the proprietor to justify.

\textsuperscript{262} Without this provision, a user who discovers that one item of data in a large database is not publicly accessible could take the entire database with impunity.
alternative sources and the user failed to use reasonable diligence in obtaining the data from those sources.

This provision imposes a significant responsibility on the database proprietor to keep records of her data sources. However, as long as the database proprietor keeps this record, she can ensure that her database is protected and that she can enforce her rights in litigation. For example, a scientist who produces a significant amount of data cannot prevent other scientists from using the data to criticize her study by insulating the data through this statute. She must either provide the data through other publicly accessible means or forego protection. Thus, Micromedex could protect Poisindex only if the poison antidotes within it were publicly accessible through alternative means. If Micromedex can prove that the antidotes were accessible, it can stop pirates from extracting its data wholesale.

In the ideal application of these suggested amendments to House Bill 354, the database user would gain the option of paying the database proprietor for the data directly or going to a publicly accessible source to obtain the data. Thus, the database proprietor is paid for the convenience offered by her database, not for her access to otherwise monopolized knowledge.

b. Resolving the Re-compilation Problem

Even if the sole-source problem is resolved, database protection will still inhibit pirates from creating new databases that combine parts of others because of the recompilation problem. The problem is that even if the data in a database is publicly accessible through alternative means, or can be reproduced in an inexpensive way, a database proprietor nonetheless can refuse to license the extraction or use of her protected database to a third party. The proprietor may very well do this where the third party is a direct competitor. Although the suggested solution to the sole-source problem should mitigate the effects of monopolistic protection of data, additional provisions are needed to ensure sufficient competition among database makers and growth through new databases.

One possible solution is to add a broad provision that precludes an apparent monopoly over data. As an example, Recital 47 of the EU Database Directive states that “in the interests of competition between suppliers of information products and services, protection by the sui generis right must not be afforded in such a way as to facilitate abuses of a dominant position . . . .” This provision invokes something like a doctrine of unfair competition to prevent unfair refusals to grant licenses on nondis-

263. Note that if a person tried to block any access to her data, it would be unlikely that her data would constitute a “collection of information,” even under House Bill 354. “Collection of information” is defined as “information that has been collected and has been organized for the purpose of bringing discrete items of information together in one place or through one source so that users may access them.” H.R. 354 § 2 (proposing 17 U.S.C. § 1401(1)). Such a person would be bringing the information together to support a study, not “so that users may access” the information. Id.

264. Database Directive, supra note 120, at recital 47. Recital 47 continues by stating that “in particular as regards the creation and distribution [sic] of new products and services which have an intellectual, documentary, technical, economic, or commercial
criminatory terms. For example, “in August 1997, the Court of Appeal of Amsterdam in the Netherlands held in *Denda International v. KPN* that a refusal by [a Dutch company] to license its ‘White Pages’ telephone directories for re-publication on CD-ROM would be an abuse of a dominant market position.”

Although such a provision is broad enough to prevent obvious monopolies, it lacks definiteness and leaves broad discretion to the courts. No rational legal adviser would rely on such an exception to evade liability. However, there is an alternative. A provision should be added to House Bill 354 requiring a compulsory license to use an otherwise protected database in certain circumstances and under definite procedures. Under the present version of House Bill 354, a person who wishes to use a protected database would have to obtain a license, the terms of which would be fixed through negotiation with the database owner. However, under a compulsory license, a person has the right to obtain a license to use the rightholder’s work without permission, provided she follows the statutory procedure and pays fixed royalties.

The arguments in favor of a compulsory license are that it would reduce the transaction costs involved in licensing works through private markets, provide a definite procedure on which users may rely to avoid liability, enhance competition, and allow for the creation of new databases by permitting new database compilers to incorporate data from otherwise protected databases. Furthermore, the criticism that compulsory licenses undermine the fundamental rights of authors does not apply in the context of database protection where the compiler is not an author.

The design of a compulsory license for databases may be informed by the four compulsory license management mechanisms contained in the Copyright Act. Generally, the two chief characteristics of any compulsory license system include provisions for (i) how the prospective licensee contracts for the license either through the Copyright Office, some similar agency, or the individual owner of the protected work, and (ii) how the royalty payments are calculated and distributed. For example, one compulsory license under the Copyright Act requires the prospective licensee of a protected work to provide the owner of the work with notice of her intention to obtain a compulsory license before distributing a copy of the work and within thirty days of actually making the copy. The licensee is...
then required to pay the rightsholder a monthly royalty fixed by statute.\footnote{For every phonorecord made and distributed on or after January 1, 1998 and before January 1, 2000, the rate is either 7.1¢ per record made based on a copyrighted recording or 1.35¢ per minute of playing time of a record based on a copyrighted recording, whichever is greater. 37 C.F.R. § 255.3 (1999). In a slightly different context, the royalty rate is fixed through voluntary negotiations or through the assessment of a copyright royalty arbitration panel.} Failure to follow the statutory procedures results in infringement. This system is amenable to a context where the protected work is taken from a single source and copied or distributed a specific number of times.

With compulsory licensing for the use of a protected database, the prospective licensee should be able to contract with the individual owner of the database because the licensee is taking the work from a single identifiable source.\footnote{One compulsory license under the Copyright Act requires the prospective commercial user who provides customers with a protected work to file specific information in the Copyright Office. This information includes its subscribers and accounts, the list of copyrighted works regularly taken, and a payment of statutorily calculated royalties to the Register of Copyright. This is the mechanism for the Cable License, which establishes a compulsory license for secondary transmissions by cable television system. 17 U.S.C. § 111 (1999). Failure to follow the statutory procedures results in infringement. The Librarian of Congress, under the recommendation of the Register, distributes the collective royalties to parties who file claims of entitlement with the Register. This system is effective where the commercial use of the licensee may be quantified in terms of accounts with buyers of the licensee's services. A related compulsory license fixes the royalty amount for a subscriber of satellite carrier television transmissions to the public for private use. See 17 U.S.C. § 119 (1999). Moreover, compulsory license is amenable to situations where the licensee is taking work from multiple rightsholders, which is dealt with by pooling the royalties from various licensees and distributing them proportionally to rightsholders. However, in the context of database protection, such a system is not necessary because the identity of the rightsholder is known.} However, determining a fair method for fixing the royalty payment is more challenging.

Various approaches for fixing the royalty amount are available. First, the price per item of data could be fixed by statute. The problem with this approach is that not all data are equally valuable. Therefore, a licensee could obtain a compulsory license for only the most valuable data, leaving the database proprietor in a disadvantaged position with her competitors. Second, the royalty could be based on the actual or potential harm to the database proprietor's market. This seems like the best indication of the real price of data in a database. However, proving harm in a dollar amount would probably prove too difficult for a database proprietor. Third, the royalty amount could be fixed by something like the copyright royalty arbitration panel, a third party through which the licensee and licensor may fairly negotiate a royalty amount. Fourth, the royalty could be fixed as a proportion of the cost to the database compiler in gathering, organizing, or maintaining the database. Because the database proprietor must maintain records of such costs to show that such expenses were substantial, as an element of a prima facie case of violation of the statute, the information required to set royalties is available and also personalized to each database
and database proprietor. Fifth, the statute could leave several options for determining the royalty amount for the plaintiff to choose, thereby allowing the party with the greatest interest in the amount, the database owner, to select the optimal approach on a case-by-case basis.\footnote{Another approach is to allow the court to determine a “reasonable royalty.” See 35 U.S.C. § 284. The amount would reflect the licensing fee the user would pay in a hypothetical negotiation with the database owner.}

Although the details remain to be worked out, a compulsory license would significantly abate the recompilation problem. Thus, a new database maker who wished to borrow from Poisindex could legally force Micromedex to license its data. This would create new opportunities for competition and effectively set a ceiling on the price that a database proprietor could demand for the use and extraction of its database contents.

c. Eliminating Freedom of Contract Issues

Another problem is that users may be forced to contract away their rights under the statutory exceptions to House Bill 354. The EU Database Directive attempts to deal with an analogous problem. It provides that “[a]ny contractual provision contrary to Articles 6(1) and 8 shall be null and void.”\footnote{Database Directive, supra note 120, art. 15; see REES & CHARLTON, supra note 82, at 172.} Article 6(1) deals with copyright protection for creative databases and Article 8 states that “[t]he maker of a database which is made available to the public in whatever manner may not prevent a lawful user of the database from extracting and/or re-utilising insubstantial parts of its contents, evaluated qualitatively and/or quantitatively, for any purposes whatsoever.”\footnote{Database Directive, supra note 120, art. 8; see REES & CHARLTON, supra note 82, at 170.} The Directive keeps parties from contracting around this express exclusion from database protection.

A similar provision should be added to House Bill 354 stating that any contractual provision contrary to Sections 1403, the “permitted acts” section, or 1404, the exclusions section, is unenforceable in an action under the statute.\footnote{It is questionable whether each of the permitted acts or exclusions should be protected under a contract-preemption provision. For example, users arguably should be able to surrender by contract those rights under § 1403(f), which provides something like a first-sale doctrine. Otherwise, a consumer, such as a library, could purchase a database and provide rental access to the database, directly harming the market for the database.} Without such a provision limiting contract law, the exceptions and exclusions could effectively dissolve under shrink-wrap\footnote{A shrink-wrap license is a license contract contained on the packaging of software. Use of the software is an acceptance of the agreement.} and click-wrap\footnote{A click-wrap license is a license contract contained on the entrance page of an Internet site. Clicking on the “acceptance” button of the page is an acceptance of the agreement.} licenses.

The proposed provision would ensure that a database provider could not force users to sign a contract surrendering their rights under the “permitted acts” and “exclusions” sections in House Bill 354. For example,
even if Micromedex provided access to Poisindex through an Internet site that required the acceptance of a click-wrap contract eliminating the user's rights under Sections 1403 and 1404, the addition of this contract-preemption provision would give Micromedex no cause of action under the statute against a user falling within the exceptions or exclusions. At the same time, Micromedex should be able to bring an action in state court for breach of contract if a user violates the terms of the agreement. A database proprietor should still be able to restrict the access and use of the contents of its database through a contract; however, she should not be able to override statutory exceptions through forced contracts. Such a provision would prevent savvy database providers from benefiting from statutory sui generis rights while circumventing the statutory restrictions on the sui generis right by forcing users to contract out of the statutory exceptions and exclusions.

d. Mitigating the Perpetual Duration Problem

The term of protection for databases under House Bill 354 is potentially perpetual because a database proprietor may mix protected and unprotected data. This problem is particularly difficult because most useful databases are dynamic, and in many instances, new data will be added to a preexisting database. There are basically three ways to deal with this dilemma.

First, the statute could state that the entire collection of data in a protected database enters the public domain after any portion of the database expires under the fifteen-year term. Thus, if Micromedex first compiled Poisindex in 2000, then even though Micromedex subsequently added a separate set of data in 2005, the entire database, including the addition, would enter the public domain in 2015. The problem with this approach is that Micromedex may be encouraged to create separate databases to ensure the new database obtains protection, even if the combination of information is not the most efficient form for the database. In other words, if old data corrupts the protection of new additions, database proprietors may be reluctant to make additions in the most efficient way.

Second, the statute could state that although an addition to or alteration of the data in a protected database can engender a new term of protection for the added portion, the database must clearly indicate which items of data are protected and which items are in the public domain. In other words, where some portions of a database may be protected while others are not, as under the current version of House Bill 354, the database proprietor must keep those portions clearly separate.

There are several problems with this approach. First, it would be costly for a database proprietor to label every item of data as either protected or unprotected. Those costs would no doubt be transferred to consumers. Second, this approach could create a thorny factual issue in litigation. Even if the database-proprietor plaintiff must show that the extracted data was protected at the time of extraction, the database proprietor could easily manufacture the required evidence since all the evidence
required to prove this element would belong to the plaintiff. It is unlikely that a user would know when the proprietor first placed the extracted information into the database.

Third, the statute could state that an addition to or alteration of the data in a protected database can extend the protection to a new term of fifteen years, but only if the addition or alteration fundamentally changes the character of the original database. This is similar, but not identical, to the predecessor of House Bill 354 and the EU Database Directive. The Directive states:

Any substantial change, evaluated qualitatively or quantitatively, to the contents of a database, including any substantial change resulting from the accumulation of successive additions, deletions or alterations, which would result in the database being considered to be a substantial new investment, evaluated qualitatively or quantitatively, shall qualify the database resulting from that investment for its own term of protection.\textsuperscript{279}

The standard for determining whether a database qualifies for a new term of protection under the Directive is whether the alteration constitutes "a substantial new investment." However, by raising the standard to whether an alteration constitutes a fundamental change in the character of the original database, a database proprietor would have a higher hurdle for obtaining new protection. Therefore, she would have to alter the fundamental character of the database every fifteen years to obtain protection.

The fundamental character of the database could be cashed out in quantitative terms where it is fundamentally changed and thereby consists of less than half of the original items of data. Thus, doubling the size of the database while keeping all of the original data would constitute a fundamental change. Deleting half of the original data would by itself constitute a fundamental change. Also, deleting twenty-five percent of the original items of data and adding a number of items equal to twenty-five percent of the original database would also constitute a fundamental change. Although qualitative factors—substantial changes in the organization of the data, substantial investments in verifying and gathering new data—could also play a role in characterizing the fundamental character of a database, quantitative factors have the virtue of definiteness.

This option is problematic because it creates the possibility of obtaining perpetual protection. However, if there is no feasible alternative, raising the standard for obtaining extended protection is the best option because it may limit the instances where a proprietor may obtain extended protection. Moreover, the potential for perpetual protection becomes less worrisome where the protection afforded to databases is weakened under new mitigating provisions for the proposed legislation, such as those suggested. All things considered, the current term provision in House Bill 354 should be replaced by a provision stating, in effect, that an addition to or alteration of a collection of information that fundamentally changes the

\textsuperscript{279} Database Directive, supra note 120, art. 10(3).
character of the original database begins a new fifteen-year term of protection for the entire database.

e. Switching the Burden of Proof

House Bill 354 places the burden of proof on the alleged infringer to prove that one of the enumerated exceptions applies and she is not liable.\textsuperscript{280} This may cause users to be overly cautious with otherwise legal uses and extractions of data. A solution to this problem is to include a provision placing the burden of proof on the database-proprietor plaintiff to prove that the alleged pirate did not fall within an exemption under the “permitted acts” section of House Bill 354.\textsuperscript{281} For example, in addition to the other elements for a prima facie case, the plaintiff would have to prove that (i) the defendant did not extract or use the information for nonprofit educational, scientific, or research purposes, or (ii) if the defendant did extract or use the information for nonprofit educational, scientific, or research purposes, the extraction or use directly harmed the actual market for the product or service.\textsuperscript{282} Placing the burden of proof on the plaintiff will make an action under the statute significantly more difficult and costly for the plaintiff and curb the plaintiff’s ever-present threat of litigation, which could otherwise stifle legal uses of protected databases.

f. Ending the Indefiniteness

The vagueness and imprecision of the terms and exceptions of House Bill 354 may cause a chilling effect in the scientific and academic communities. The problematic expressions include “substantial part,” “harm to the actual or potential market,” and “through the investment of substantial monetary or other resources,” each of which occur in proposed Section 1402.\textsuperscript{283} Moreover, the blanket exception for the use or extraction for “educational, scientific, research, and additional reasonable uses” is too indefinite. The following considers possible provisional cures for each problem.

1) Defining Substantial Part

Two provisions should be added to House Bill 354 to deal with the vagueness of “a substantial part, measured either quantitatively or qualitatively,” which is the standard used to determine whether the defendant misappropriated a sufficient amount of the database to incur liability.\textsuperscript{284} The notion of a “substantial part” only makes sense relative to a particular


\textsuperscript{281} Id. (proposing 17 U.S.C. § 1403). Note that instead of requiring the plaintiff to meet a burden of proof, one could require that the plaintiff have the burden of production—a burden to produce some evidence on that issue to avoid an adverse decision. Once some evidence is brought forth, the burden could switch back to the defendant to prove its truth or falsity.

\textsuperscript{282} Id. (proposing 17 U.S.C. § 1403(a)(1)).

\textsuperscript{283} Id. (proposing 17 U.S.C. § 1402).

\textsuperscript{284} Id.
database; a substantial part of a small database may consist of a few items of data whereas a substantial part of a large database may consist of many items of data. Therefore, both “substantial part” and “collection of information” need to be made more precise.

First, the collection of information should be fixed to some extent by the constitution of the database at the time of the alleged misappropriation. Thus, if Micromedex has 10,000 poison antidotes at the time a pet store extracts 2000 and Micromedex subsequently eliminates 5000 items from its database, the working collection of information for the purpose of this issue is the original 10,000 poison antidotes. Second, the definition of “collection of information” should concede the possibility of there being a database within a database. For example, a database of real estate listings throughout the United States may consist of fifty independent databases, one for each state. Thus, where an alleged pirate extracts all the listings for Nebraska, this amount should, at least conceivably, be measured against the database for that state rather than the database containing each state combined. The difference is one of extracting the whole database or a very small portion—perhaps less than one-fiftieth—of the overall database. Otherwise, large databases would be inherently disadvantaged under the statute. Third, restrictions should be placed on the database proprietor’s ability to define her database on an ad hoc basis. If the database proprietor is able to define her database any way she pleases, she can always prove that an alleged pirate extracted or used the whole of that arbitrarily defined database.

Delineating useful restrictions is no easy task. One consideration in restricting the definition of a database is whether that collection of information would have an independently sustainable market. For example, if there is an independent market for a collection of Nebraska real estate listings, then the proprietor could conceivably define her database as a compilation of Nebraska listings. This may also allow databases for large cities such as Chicago and New York, while disallowing a proposed database for Ithaca, New York listings, depending on the market demand for such information. Another factor that could restrict the definition of a database is the separability of the information. Some types of data are only valuable as a collection. As such, these databases should be defined as the larger set of items. A third factor is the searchability or categorization of the information. For example, if Poisindex is not categorized according to animal-poison antidotes and the database cannot be searched to produce a listing of animal-poison antidotes, then Micromedex should be precluded from claiming that the collection of animal-poison antidotes

286. Credit for this idea should be given to Peter Martin, Professor of Law at Cornell Law School.
287. Credit for this idea should be given to Jeffrey Sickel, a computer programmer in Chicago, Illinois. He noted that some market research databases are only valuable as a whole. Individual items have no value because only the collection of data as a whole provides the required analysis for market research.
is an independent database.\textsuperscript{288}

The notion of a "substantial part, measured either quantitatively or qualitatively" should also be defined to give users a better idea of how much data they can take from a protected database without incurring liability. Although House Bill 354 explicitly permits the extraction of individual items of data, the statute, or additional regulations, should set out quantitative guidelines. For example, a guideline could state that any individual may extract or use up to five percent of an overall collection of information, measured against the database existing at the time of extraction,\textsuperscript{289} over a period of five years. Thus, if Poisindex contained 10,000 poison antidotes as a whole, a person could extract up to 500 of the items over a period of five years without the risk of incurring liability.\textsuperscript{289}

Fixing the actual percentage amount, whether at five percent or some other level, should provide a compromise. The fact that most databases are valuable because they are comprehensive suggests that taking only five percent of the items from a protected database will not significantly hurt the original database in any significant way. However, if the items in a database are independently valuable and a pirate is capable of selectively taking only the most valuable items in the database, she could undermine the database proprietor's investment by selecting and extracting the most valuable five percent. Thus, a specific percentage should be chosen that balances these and other considerations.

Under this construction of "substantial part," a user may extract or use fixed portions of databases without the chance of infringing the database owner's statutory rights. However, it should not be presumed under the statute that extracting more than a specific percentage constitutes infringement. Rather, the amount simply more specifically defines

\textsuperscript{288} Although this factor will be dependent on somewhat arbitrary categorizations of data and electronic databases with powerful engine software would be advantaged, it could preclude database makers from completely arbitrary definitions of their database.

\textsuperscript{289} The clear exception for a certain quantitative amount probably should be measured against the overall database, as opposed to a database within a database, since the user will probably not be able to estimate how the proprietor will try to define her working database at the time of litigation. A database within a database is not clearly defined and would not provide sufficient definiteness.

This does however allow a problem where a user extracts a few items of data in 1999, when the database contained, one hundred items, but then in 2000, several thousand items are added to the database. Even though in 2000, five percent of the current database may consist of over one hundred items, the user would be restricted to extracting only five items. In other words, the first extraction fixes the total number of allowable extractions. This problem is slightly tempered by the fact that a database proprietor could conceivable eliminate items of data and create a situation where the user could legally extract the whole database under the five-percent formula.

\textsuperscript{290} If such a provision is added to database legislation, the statute may also have to impose certain formalities upon database proprietors. For example, the statute probably should require that a database owner publish the total number of items of data in the database at any particular time. Thus, Micromedex would always provide users with the total number of poison antidotes in their Poisindex, allowing the user to decide how many items he can legally extract.
exempt amounts for the user.\textsuperscript{291}

In summary, provisions should be added to House Bill 354. One provision should state that "collection of information" in the context of a use or extraction of a substantial part means a collection of items defined in terms of the ability for the collection to sustain an independent market, the separability of the information, the categorization or searchability of the overall collection, and other relevant factors at the time of the alleged misappropriation. Another provision should define "substantial part" as more than, say, five percent over a five year period of the overall collection of information at the time of the first extraction or use. Beyond that amount, a substantial part is measured relative to a particular "collection of information" and measured quantitatively and qualitatively.

2) Defining Harm to the Market

A provision should be added to House Bill 354 to sharpen and explicate the expression "harm to the actual or potential market."\textsuperscript{292} Although the statute needs further clarification, it contained some additional language describing the type of harm contemplated. First, the statute speaks to "[a]ny person who extracts, or uses in commerce" a proscribed amount of the database and further states that the database proprietor only obtains protection for her database if it "is offered or intended to be offered for sale or otherwise in commerce."\textsuperscript{293} These restrictions are required for constitutional reasons, since the source of the statute would be the Commerce Clause, not the Intellectual Property Clause.\textsuperscript{294} The first restriction implies that a noncommercial use, although not an extraction,\textsuperscript{295} for private purposes may not constitute an infringement. Thus, the bill explicitly allowed any personal use of a protected database. The second restriction implies that a database proprietor cannot protect her database unless she sells or intends to sell her database.\textsuperscript{296} Thus, a database owner cannot sit in the slack.

\textsuperscript{291} In addition, some qualitative elements of a "substantial part" may add specificity, although such additions are by their nature less definite than a quantitative measurement.


\textsuperscript{293} Id.

\textsuperscript{294} Section 1401(4) defines "commerce" as "all commerce which may be lawfully regulated by the Congress." Id. (proposing 17 U.S.C. § 1401(4)).

\textsuperscript{295} House Bill 354 states that "[a]ny person who extracts, or uses in commerce, all or a substantial part . . ." of an otherwise protected database is potentially liable. As written, it does not limit extractions to commercial extractions. Id. (proposing 17 U.S.C. § 1402). However, it is not clear whether the bill should include all extractions. For example, suppose a person extracts the entirety of a protected paper version of a directory to his computer for his own personal convenience and never provided others with access to his electronic copy. Presumably, he would not be violating the statute even though he harms the owner's actual or potential market in the strictest sense. However, because the language proscribes any extractions, not just commercial extractions, such a person could be held liable. Arguably, such a person should be prevented from personal extractions.

\textsuperscript{296} This is significantly different from copyright law. Under copyright law, a work can still be infringed upon even where the author or owner of the right does not sell or intend to sell the work. Moreover, because of this divergence from copyright law, showing harm to the market is more difficult in the case of databases. Under the fair use
on her data and obtain protection for her database. This language significantly confined the analysis of a possible violation to commercial uses of a commercial database.

Second, “potential market” is defined under House Bill 354 as “any market that a person claiming protection under Section 1402 has current and demonstratable plans to exploit or that is commonly exploited by persons offering similar products or services incorporating collections of information.” This limits the scope of potential markets. However, both this language and the language discussed above fail to provide any guidance as to the amount or type of harm needed to trigger a violation of House Bill 354.

Obviously, actual harm to the database proprietor is strong evidence against an alleged pirate. For example, if Micromedex could show that it lost a contract because a potential customer obtained the information she sought from a pirate, then harm is shown dispositively. Since this type of evidence would probably be rare, a provision should be added defining “harm” as demonstrable harm, probable harm, or the meaningful likelihood of future harm.

Borrowing from copyright law, “probable” and “meaningful likelihood” of harm could be cashed out in terms of factors such as whether the use diminishes the potential sale or license of an otherwise protected database, whether the use tends to interfere with the database’s marketability, or whether the use fulfills the demand for the original database. Common to each of these factors are several additional factors. First, where a use or extraction is made accessible to the public, or at least potential customers of the protected database, the use or extraction is more likely to harm the actual or potential market of the database owner. Second, where a use or extraction is made from a database for which there is a demonstrably high demand, the use or extraction is more likely to harm the actual or potential market. Third, if the original database is marketed to a limited geographical area, either intrinsically or by choice, use or extraction outside that area is less likely to harm the actual or potential market.

---

297. This precludes people from accumulating data, protecting the data, and then preventing others from doing the same through their protection. This phenomenon is more common in patent law where public interest groups patent a new invention without the intention to market it or sell it to prevent a company from obtaining a monopoly over the invention.

298. H.R. 354 § 2 (proposing 17 U.S.C. § 1401(3)).

299. Leaffer, supra note 10, at 440.

300. For example, a phone or business directory for people and businesses in Ithaca, New York is intrinsically more valuable to people in Ithaca.
Fourth, evidence that the alleged pirate made money from the use or extraction should be strong evidence that the pirate harmed the database owner's actual or potential market. For example, if an alleged pirate simply extracts Poisindex onto her computer's hard drive at no expense to Micromedex, Micromedex would have a difficult time showing that the use harmed its market where the use is private. However, if the alleged pirate's use allows her to avoid paying a license fee to Micromedex, Micromedex would have a stronger case, although still weak because the use is not commercial. However, if the alleged pirate publishes an extracted portion on her private web site, again, Micromedex would have a stronger case, although still weak because the extraction is not made widely and publicly accessible. But if the alleged pirate advertised the site and linked her site to commercial sites, Micromedex would have a strong case that her use is harming their actual and potential market. Furthermore, if the original, protected database is a successful service for which there is a high demand, the case is even stronger. The result is a continuum of actual or possible harm to actual or potential markets.

In summary, the expression "harm to the actual or potential market" can be sharpened to encompass demonstratable harm, probable harm, or the meaningful likelihood of future harm, measured against such factors as the accessibility of the pirated data, the demand for the data, whether there are geographical restraints of the marketability of the data, and whether the pirate actually made money from the use or extraction.

3) Defining a Substantial Investment to Trigger Protection

Provisions should be added to House Bill 354 to explain the expression "gathered, organized, or maintained by another person through the investment of substantial monetary or other resources." It is clear from the language of House Bill 354 that wholesale copying of data from another source, such as downloading data from an Internet-accessible database, would not constitute a sufficiently substantial investment to warrant protection. Depending on how the expression is interpreted, this could be a substantial bar to protection for database makers. However, unless the notions in this expression are made more definite, no user of a database dare rely on it to use or extract data from an otherwise protected database.

The notions that need to be made more precise are of two types. The expressions "gathered," "organized," and "maintained" involve the behavior for which a substantial investment can produce protection. The expression "investment of substantial monetary or other sources" involves the amount of investment and its measurement.

First, "gathered" should be defined to include collecting data from original sources and independent research. It also should include collecting data from otherwise protected sources, either legally under an exception or through licensing. Clearly, the final statute should reward the

“sweat of brow” associated with producing an original source of the data—for example, collecting addresses and telephone numbers for a new directory. Moreover, the statute should protect the costs of performing costly research that produces data. The statute, however, should also create protection if the database maker incurs sufficient costs from bringing together data from multiple sources, recompilation, regardless of whether the sources are already accessible and protected.\footnote{302}{Many current Internet search engines are databases where the data are collected by “robots” that index sites and then tabulate the results into a database that may be searched by a browser. It is not clear whether this type of “gathering” is contemplated under House Bill 354. The exclusion for “digital on-line communications” suggests, at first blush, that these databases are not protected. \textit{Id.} (proposing 17 U.S.C. § 1404(c)).}

Second, “organized” should be defined to include sufficiently substantial investments for categorizing data, obtaining and implementing search engines for electronic data, and presenting data. The EU Database Directive explicitly mentions that a database maker may obtain protection for costs incurred in “presenting” data in a database.\footnote{303}{The Directive protects any investment in obtaining, verifying, or presenting data. Database Directive, supra note 120, art. 8, § 1 recital 40.}

Third, while “maintaining” should be defined to include verifying the accuracy of the data, it should exclude substantial investment incurred in providing legal protection for the database. Thus, if Micromedex copied its original version of Poisindex from an unprotected source that did not validate the accuracy of the data, Micromedex should be able to obtain protection if they confirm and verify its accuracy. Not only does the EU Database Directive specifically mention that “verification” may engender protection\footnote{304}{\textit{Id.}} but also this protection is justified. Since Micromedex could potentially be held liable for the inaccuracy of the data,\footnote{305}{See supra Part III.A.1.b.} it should be remunerated for verifying its accuracy. At the same time, if the accuracy of the original source was already significantly validated, then independent verification by Micromedex should not engender new protection. Moreover, the costs incurred by a database proprietor in protecting her investment through legal (contractual), technological (encryption technology), or other means (general legal fees) should not count toward the maintenance of the database.

Fourth, “investment of substantial monetary or other sources” should be defined to allow for cumulative investments over a period of years, and it should specify a lower limit required to obtain any protection. The investment requirement should not discriminate between investments made completely at inception and investments made over the entire period until an alleged misappropriation. The relevant investment includes all the expenses incurred up to the point when the alleged pirate committed the questionable use or extraction. Additionally, a provision should set a minimum monetary investment requirement for obtaining protection. Since only commercial databases are protected under House Bill 354, it is reasonable to expect the database proprietor to have expenditures associated
with the database. Moreover, all the types of expenses associated with producing a commercial database have monetary equivalents; whether the resource utilized to make the database is capital, labor, or technology, the expense can be measured in dollars.\textsuperscript{306} Although the minimum investment amount should be relatively low, and perhaps increased annually based on an inflationary scale, meeting the requirement should not constitute prima facie protection. In other words, while the database maker must meet the amount for obtaining the possibility of protection, meeting the amount would not ensure protection.\textsuperscript{307}

In sum, the expression “gathered, organized, or maintained by another person through the investment of substantial monetary or other resources” can be made more definite by defining each of the substantive terms: “gathered,” “organized,” “maintained,” and “investment of substantial monetary or other resources.” Of particular importance, “maintained” should be defined to encompass verification of data, and “investment of substantial monetary or other resources” should be defined to set a minimum dollar amount of investment required for any protection.

4) Sharpening the Reasonable Use Exceptions

Regulations or guidelines should accompany House Bill 354 to explicate the blanket exception for individual acts of use or extraction for “educational, scientific, research, and additional reasonable uses.” As written, the exception contemplates a balancing test consisting of factors such as whether the use or extraction is commercial, whether the alleged pirate used or extracted the data in good faith, whether the extracted data is incorporated into a database that is similar to the original, and whether the alleged pirate is in the same, or similar, business as the original database maker.\textsuperscript{308} A bad-faith, commercial extraction by a person in the database industry who incorporates the data into a similar database as the original would not benefit from the exception. However, a good-faith, nonprofit extraction by a scientist who uses the data in a completely different context would presumably be shielded from liability. The test provides some indication of proscribed and allowable uses and extractions; however, without

\textsuperscript{306} According to one commentator, an investment, measured qualitatively or quantitatively, as used in an analogous context by the EU Database Directive includes investments through financial, human or technical resources. \textsc{Rees & Charlton, supra} note 82, at 59-60; \textit{see also} Copyright and Rights in Databases Regulations, SI (1997) 1997/3032 § 13(1) (Eng.). Under Regulation 12(1), “substantial” as measured in relation to any investment, extraction, or re-utilization means substantial in terms of quantity or quality, or a combination of both. This would include advancing capital to fund the making or renewal of a database (financial resources), hiring staff or contractors or making them available for the work on the database (human resources), and the provision of computing capacity or know-how (technical resources).

\textsuperscript{307} Conceivably, the dollar amount could be relativized to the type of behavior under which the proprietor claims protection: gathering, organizing, or maintaining. However, the database maker is being remunerated for her investment in producing a useful database, regardless of how she goes about making the improvement or creation.

\textsuperscript{308} \textit{See supra} Part III.B.2.f.
more definiteness, potentially useful and legal database uses may be avoided.

In the legislative history of the Copyright Act's "fair use" provision, a set of guidelines are laid out for classroom uses of copyrighted works. The guidelines contain a provision for single copying by teachers,\textsuperscript{309} but more appropriate to database legislation, the guidelines contain a provision dealing with multiple copies for classroom use. The gist of the latter provision is that copying is legal provided it meets (i) a brevity requirement that is defined for four kinds of works—poetry, prose, illustrations, and special works; (ii) a spontaneity requirement which is defined in terms of the possibility that the user had time to obtain permission from the copyright holder; (iii) a cumulative effect test, which specifically precludes more than "nine instances of such multiple copying for one course during one class term"; (iv) a requirement that each copy contains notice of copyright; and (v) a requirement that the use does not replace the market for the original work and the teacher does not charge her students for the copies.\textsuperscript{310} Although failure to meet these requirements does not constitute copyright infringement, if a person meets the requirements, he is almost certain to obtain an exemption from liability. These guidelines provide a useful model for research and educational uses and extractions for databases.

A similar guideline should be added to House Bill 354 to expressly exempt certain educational, scientific, and research uses or extractions of data. In particular, the guidelines should specify that a use or extraction is exempt provided it meets (i) a brevity requirement,\textsuperscript{311} (ii) a spontaneity requirement, (iii) a cumulative effect test, (iv) a requirement that the use or extraction mentions the original protected database, and (v) a requirement that the use or extraction is not-for-profit and does not replace the market for the original database. The details of the guidelines are left to be worked out; however, a guideline of this form could significantly reduce any "chilling effect" caused by the indefinite balancing test under the House Bill 354 proposal.

In sum, numerous terms in House Bill 354 should be made more precise. Although the language of House Bill 354 would no doubt become more definite as the courts develop case law interpreting the statute, the potential danger for overprotection of data requires definiteness now.

In general, a number of additional provisions and guidelines should be added to House Bill 354 to eliminate or mitigate the problems associated with the bill. In particular, provisions should be added (i) creating an element for infringement that targets proprietors attempting to monopolize the sole source of data; (ii) creating a compulsory license in certain cir-

\textsuperscript{310} Id.
\textsuperscript{311} This requirement would be similar to the limitation suggested for measuring a "substantial part"; however, in this case, the amount would be significantly more generous. For example, if a pirate must use or extract more than five percent to be liable, then a research user under the exception would specify that the person must use or extract more than fifteen percent to be liable.
cumstances; (iii) creating a limited preemption of contracts; (iv) raising the standard for obtaining a new term of protection; (v) switching the burden of proof to the plaintiff to show that the defendant does not have an exempt status; and (vi) creating more definite rules by sharpening terms in the proposed statute and specifying guidelines exempting certain educational, scientific, and research uses or extractions. By incorporating these suggestions into a successor to House Bill 354, database legislation would bear the following characteristics, where the italicized parts denote suggested changes.

---

Compromise Proposal for U.S. Database Protection

**Source**
The proposal would expand the Copyright Act with a Chapter called "Misappropriation of Collections of Information" and protect non-creative databases. However, the source of protection would stem from the Commerce Clause, not the Intellectual Property Clause of the U.S. Constitution.

**Scope**
Protection would apply to (1) factual databases, as defined by the statute; (2) for which the data is publicly accessible through alternative means, (3) gathered, organized, or maintained through an investment of substantial monetary or other resources, as defined by the statute; and (4) used in commerce.

**Rights**
The statute would prohibit (1) the extraction of all or a substantial part, as defined by the statute, of the contents of the database that (2) caused harm to the actual or potential market of the owner, as defined by the statute (3) where the alleged infringer does not have an exemption under certain specified exceptions.

**Exceptions**
The statute would offer at least the following exceptions: (1) fair use measures for scientific, education and research purposes; and (2) individuals acts of use or extraction of data for the purpose of illustration, explanation, example, comment, criticism, teaching, research, or analysis—as explicated by guidelines. The statute would also offer exclusions for government collections and news reporting. The statute would provide for a procedure through which a user could obtain a compulsory license to use or extract data from an otherwise protected database. In addition, the statute would void any contractual provision contrary to certain specified exceptions.

**Duration**
The property right in the database would continue for a fifteen-year period beginning when the database was injected into commerce, but this period could restart upon a change in the fundamental character of the original collection of information.

**Remedies**
The bill prescribed civil remedies—injunctive relief; impoundment; and monetary relief, including damages and attorney's fees—and criminal penalties including fines ranging from $250,000 to $500,000 and imprisonment for five to ten years.

The added provisions, definitions, and guidelines are needed to provide the balance between promoting investment in databases and protecting the public domain. Moreover, the compromise proposal moves toward securing international protection for U.S. databases.

2. **International Database Protection**

The compromise proposal would offer opportunities to secure international protection for U.S. databases. Not only would the compromise engender a fair chance to obtain reciprocal protection in the European Union but also it could be used as a platform for negotiating further international protection through amendments to the Berne Convention, the WIPO Treaty, TRIPS, or an entirely new international instrument.
First, although the compromise proposal contains surface differences from the requirements set forth in the EU Database Directive, the compromise has a decent chance for obtaining reciprocal protection in the European Union. To obtain reciprocal protection from EU member countries, the United States must provide “comparable protection” to the EU sui generis right as determined by the EU Council. On each of the substantive provisions in the Directive, the compromise proposal contains analogous provisions. Indeed, the EU Database Directive elaborated many of the solutions in the compromise proposal. Furthermore, even the more significant disparities between the compromise proposal and the Directive may be overlooked. Although the European Union currently benefits from the lack of protection of U.S. databases in the European Union, the United States still possesses significant bargaining power because it is the single largest market for databases and database services. Just as U.S. database makers cannot protect their databases in the European Union, under current U.S. law, EU database makers cannot protect their databases in the United States. Finally, in considering the options available for U.S. database legislation, the compromise, or some refined version of it, is no doubt the most agreeable. House Bill 1858 would not provide any opportunity for international protection whereas House Bill 354 would provide too much protection for databases in the United States. The United States has an interest in providing no more protection to foreign databases than can be achieved in the foreign jurisdictions.

Second, the compromise proposal could provide a standard for increasing international protection of databases under multilateral treaties—the Berne Convention and the WIPO Treaty—and trade agreements (TRIPS). Once the United States adopts domestic database legislation, they can negotiate added international protection for U.S. databases through an amendment to one of the existing instruments or through the creation of an entirely new instrument. Instead of resisting any international database protection, the United States could then play an active role designing protection that will favor its interests. Until the United States commits to some form of domestic database protection, they cannot play a positive role in negotiating international protection.

In sum, the compromise proposal could have the immediate benefit of obtaining reciprocal database protection in EU member countries, and it could also give the United States an edge in negotiating broader international protection through treaties and trade agreements.

Conclusion

Database legislation should accomplish three goals. It should create an incentive for U.S. businesses to invest in databases, ensure that scientists and the public continue to profit from easy access to information, and secure international protection. Although none of the contemplated legislative proposals satisfy these three goals, the compromise proposal might.

312. Database Directive, supra note 120, art. 11 recital 56.
House Bill 1258 fails to fulfill two of the goals. As written, the bill simply provides too little protection to create a sufficient incentive for U.S. businesses to invest in databases. Moreover, the bill fails to secure needed international protection. Under the current international framework, U.S. database makers cannot achieve any protection for non-creative databases unless Congress enacts legislation comparable to the EU Database Directive. House Bill 1258 does not provide comparable protection. For both of these reasons, it cannot resolve the database dilemma.

House Bill 354 fails to fulfill the other remaining goal; it does not ensure that scientists and the public continue to profit from access to information. Although the sui generis right under House Bill 354 would likely secure reciprocal international protection under the EU Database Directive, it unnecessarily provides too much protection for database makers in the United States. For this reason, it cannot resolve the database dilemma.

Compared to House Bill 1858 and House Bill 354, the compromise proposal comes closest to meeting each challenge. While going far beyond House Bill 1858 in creating an incentive for U.S. database makers, it avoids the overly protective regime of House Bill 354. Moreover, the compromise proposal is designed with an eye toward international protection and thus offers a fair chance for gaining some international protection under the EU Database Directive. All things considered, the compromise proposal comes closest to resolving the database dilemma in the United States.