

TESTIMONY TO THE  
SUBCOMMITTEE ON COURTS AND INTELLECTUAL PROPERTY  
COMMITTEE ON THE JUDICIARY  
U.S. HOUSE OF REPRESENTATIVES

LEGISLATIVE HEARING  
ON  
H.R. 354, THE "COLLECTIONS OF INFORMATION ANTIPIRACY ACT"

MARCH 18, 1999

PRESENTED BY:  
CHARLES E. PHELPS  
PROVOST  
UNIVERSITY OF ROCHESTER

ON BEHALF OF:  
Association of American Universities  
American Council on Education  
National Association of State Universities and Land-Grant Colleges

---

I am Charles E. Phelps, Provost of the University of Rochester. I appreciate this opportunity to testify before the Subcommittee on H.R. 354, "The Collections of Information Antipiracy Act." My testimony is presented on behalf of the Association of American Universities, the American Council on Education, and the National Association of State Universities and Land-Grant Colleges, which together represent over 1,500 colleges and universities. These Associations understand the need to protect databases, and they support legislation targeted to address unfair competition and database piracy. Indeed, universities and colleges often are creators of collections of information.

We very much appreciate the revisions to H.R. 2652 contained in H.R. 354 that seek to address concerns raised about acceptable uses of databases and the potential for perpetual protection of databases. We are concerned, however, that the protections provided to collections of information in H.R. 354 remain overly broad in a number of key respects that will impede the core academic activities of research and teaching. We are prepared to work with the Subcommittee and with other interested parties to develop a consensus approach that can be supported by all.

We approach the issue of database protection with a single core principle: Because data and information are the cornerstone of scientific and scholarly research, teaching and learning, we believe it is imperative to preserve the fundamental premise of this nation's information policy that no one may own facts or information, or may prevent the full, unfettered use of facts and information. As the Supreme Court said in *Feist*, "all facts - scientific, historical, biographical, and news of the day . . . are part of the public domain available to every person." *Feist Pubs., Inc. v. Rural Telephone Service Co.*, 499 U.S. 340, 348 (1991), quoting *Miller v. Universal City Studios, Inc.*, 650 F.2d 1365, 1368 (5th Cir. 1981). "[T]he raw facts [in a compilation] may be copied at will. This result is neither unfair nor unfortunate. It is the means by which copyright advances the progress of science and art." 499 U.S. 340, 350 (1991).

This policy has served the country well. The United States stands at the forefront of learning, science and technological achievement, and the nation has benefited richly from this leadership in international

economic competitiveness, lifesaving advances in medicine and health care, technological superiority in defense, and an enriched quality of life for our citizens. We believe that the enlightened information policies of this nation have played a significant role in sustaining the creativity and productivity of the research and education programs that led to these benefits. Congress should avoid any legislation that could threaten this fundamental principle that facts and information remain in the public domain. Educators and researchers should not be required to have an attorney on call at all times. Wherever decisions are to be made about the proper scope of protection for compilations of information, Congress should err on the side of caution and access to information.

Based on this principle, we can identify several critical standards that any legislation to protect compilations of information should meet: First, protection should be targeted to deal with specifically identified wrongful conduct. Second, protection should be addressed to a clearly defined class of materials and should be limited to compilations as *compilations*, not the facts or the information per se. Third, the protection available for compilations of information should be no broader, or stronger, than the protection available for original, creative works of authorship.

Regrettably, H.R. 354 fails to meet these standards. We do not lightly conclude that, as written, H.R. 354 threatens to chill research and education and to place information in the control of a limited number of commercial interests. But we do conclude that we cannot support H.R. 354 in its current form.

Our three associations and their member colleges and universities have the following specific concerns.

- H.R. 354's prohibition on "extraction or use in commerce" is unreasonably broad and grants substantial control over information itself, long after it is extracted from a protected collection. The bill should focus its prohibition on one who makes extracted material available to others in a manner that is likely to serve as a market substitute for the original collection. Certain extraction may also be prohibited.
- The definition of a protected collection is so broad that it literally covers almost any publication, including an article, a textbook, or the report of a scientific study (with accompanying data). It should be narrowed in the text of the bill.
- Liability under H.R. 354 should not be triggered where a taking is quantitatively or qualitatively insubstantial or where the portion that is taken was not the subject of substantial investment.
- The bill's concept of "market harm" far exceeds the traditional bounds of misappropriation and unfair competition law and provides database proprietors with the ability to create markets, and thus, liability.
- The exception for non-profit educational activities contains a broad, vague condition that vitiates its protection.
- The exception for other reasonable uses is insufficiently flexible, and contains conditions that greatly limit its benefit.
- The bill fails to protect the public against unreasonable market power from compilations that are not readily available from competitive sources.
- The bill should secure access to older versions of protected compilations, in order to prevent perpetual protection.
- The bill does not protect institutions that act as online service providers from unreasonable liability.

- The bill lacks clear exemptions from liability for non-profit teaching activities, akin to the exemptions in Copyright Act section 110.

In the following statement, the Higher Education Associations provide an overview of the basic academic activities that are threatened by H.R. 354. We then amplify the three standards we have identified. Finally, we discuss our specific concerns with the bill and offer suggestions for improving the legislation. It is important to stress that these suggestions are not presented as a menu from which a few items may be chosen. We believe that each of the fundamental issues discussed below should be addressed in order to preserve the flow of information and the progress of science and learning.

## **I. The Academic Environment and Activities Threatened by H.R. 354**

The research and teaching missions of colleges and universities are fundamentally tied to information and the translation of information into knowledge: through the production, analysis, verification, interpretation, and dissemination of information, scientists and scholars expand the frontiers of knowledge and transmit that ever-expanding knowledge to colleagues and to students. The results of research are publicly disseminated through articles, books, workshops, conferences, and increasingly through digital networks as well. Research results so disseminated are used by other scientists and scholars - to build on, to critique, to re-examine and re-interpret. Through the give and take over what may be initially conflicting data or interpretations of data, new phenomena are understood and verified, and knowledge is advanced.

The process of translating data into knowledge requires the open exchange of information among allied scholars and critics alike. Increasingly, research is conducted in teams, often from several institutions. Data are drawn from multiple sources, recombined and merged with new data to produce data sets that may lead to new and unanticipated findings. Data sets vary from the results of a single experiment, captured in a table in a single journal article, to the vast databases of information compiled from meteorological remote sensing instruments, geographic information systems, particle accelerators, and systematic aggregations of research results to produce databases of genomic, chemical, and medical information, and much more.

Databases supporting research and scholarship are not limited to the sciences. Databases supporting work in the humanities and social sciences are proving increasingly essential to advancing knowledge in these disciplines; specialized dictionaries, annotated bibliographies of worldwide research resources, census information, and compilations of text citations are just a few of the systematic compilations of information critical to humanistic and social science research.

In the academic community, these databases are dynamic instruments: they are not only sources of information, but they themselves - or components of them - become ingredients in new products, both through the combination of multiple contemporaneous data sets to produce qualitatively new products, and through the re-analysis of prior data from new perspectives provided by new findings or new analytic tools. A scientist may apply a formula developed from his or her research to a different set of data, yielding a different interpretation of that data; multidisciplinary researchers may combine components of databases from physical, biological, chemical, and meteorological data to understand the dynamics of ecological systems; social scientists may combine elements of databases of demographic, economic, legal, and political information in comparative analyses of national or regional populations worldwide.

Some of the best education is learning by doing and by discovering, and students are increasingly using databases to draw their own conclusions, duplicating the research process to learn through discovery under the guidance of faculty.

For all of these research and educational activities, faculty and students must be able to have open and easy access to compilations of data of all sizes, from single research results to large databases, and

they must be able to work with these compilations - extracting, combining, and aggregating sets of data - to advance the frontiers of knowledge and educate students about those advances.

These academic uses of information do not require that all information be free; indeed, universities now pay substantial sums for commercial databases. But these uses do require sufficiently flexible conditions of use, conditions that can be stultified by a proprietary protection scheme that makes use, reuse, and recombination difficult and militates against the ability to exchange information with colleagues and students.

## **II. The Standards Against Which Legislation To Protect Compilations Should Be Judged**

In general, the Associations share the view of the Administration, as expressed last year by the Department of Commerce, that "any [law to protect compilations and databases] should be predictable, simple, minimal, transparent, and based on rough consensus." Letter from Andrew J. Pincus, General Counsel, Department of Commerce, to Senator Patrick J. Leahy, August 4, 1998 (the "Pincus Letter"). In particular, we emphasize three important criteria.

First, the protection should be targeted to deal with specifically identified wrongful conduct. We respectfully believe that it is not good policy to adopt a broad, catch-all prohibition, which is subject to potentially broad but ambiguous exceptions that are subject to judicial construction and application. Such an approach will ensure that any activity that arguably falls within the scope of proscribed conduct will need to be evaluated by attorneys. Researchers, educators and scientists perform functions that serve vital public purposes. They should not be required to have an attorney looking over their shoulder at all times. I can assure you that, while this might be good for our attorneys, it will not be good for scholarship or for science.

Second, protection should be addressed to a clearly defined class of materials. If the goal is to protect incentives for the creation of large databases that require extensive effort to develop and organize, the legislation should be crafted to apply to just such works. The risk of spill-over into other types of works should be minimized. Further, it is essential that the legislation protect the compilations as *compilations*, not the facts or the information contained in the compilations *per se*. While this is a difficult line to draw, it is critical that it be drawn properly.

Third, in no respect should the protection available for compilations of information be broader, or stronger, than the protection available for original, creative works of authorship. Similarly, the exceptions and privileges applicable to the new legislation should be no narrower in any respect than the exceptions and privileges applicable to copyrighted works. Given the importance of the free flow of information to learning and science, there can be no justification for granting broader protection for compilations than exists for creative, copyrighted works.

## **III. Fundamental Concerns With H.R. 354**

### **A. The Legislation Should Not Broadly Target the "Use" of Information.**

The conduct proscribed by the H.R. 354's operative prohibition, "extraction or use in commerce" is unreasonably broad and creates a danger that database proprietors will be permitted to exercise substantial control over information itself, long after it is extracted from their protected collection. By including "use" as a prohibited act, H.R. 354 violates our first and third standards.

The term "use" has no meaningful bound. Copyright law does not even prohibit "use;" rather it prohibits five defined acts (reproduction, public distribution, creation of derivative works, public performance and public display). 17 U.S.C. section 106. "Use" conceivably covers many acts fully permitted by copyright law (including reading, research, lecturing about, discussing, using to support

debate, etc.). Indeed, any activity involving information from a collection is a form of "use."<sup>1</sup> At each step of the way, a professor or student would be required to know, on pain of liability (i) whether the information originated in a protected collection of information, (ii) whether the collection was gathered, organized, or maintained by another person through the investment of substantial monetary or other resources, (iii) whether the information used represents a quantitatively or qualitatively substantial part of collection from which the information originated, (iv) whether the specific use is the subject of a license to use the information, (v) the actual and potential markets for the collection of information, (vi) whether the specific use will cause harm to one of the actual or potential markets, and (vii) whether the conduct falls within one of the several exceptions to the rights granted by the legislation. Many of these questions will require knowledge of unknowable facts. Many will require knowledge and understanding of ever-developing judicial interpretation. This simply makes no sense. It is a boon to lawyers. It will be a bane to scholarship, science and education.

***Proposed Alternative.*** We submit that database legislation should be targeted at specific wrongful conduct. We understand and accept the need to limit the ability of competitors or others who would offer a true substitute collection from free riding on the work of the originator. This purpose may be served by creating a cause of action against one who extracts substantial portions of a collection and makes the extracted material available to others in a manner that is likely to serve as a market substitute for the original collection.<sup>2</sup> This is comparable to the approach recommended last year by the Administration - "there should be effective legal remedies against 'free-riders' who take databases gathered by others at considerable expenses and reintroduce them into commerce as their own." Pincus Letter at 1. It also is comparable to the approach taken by the alternative "Fair Database Competition" bill placed into the Congressional Record for discussion purposes by Senator Hatch on January 19, 1999.<sup>3</sup> The central focus of any legislation should be to prevent free-riding competition in the marketplace.

We also understand that there is a desire simply to prohibit unauthorized extraction of information from protected collections. The concern, as we understand it, is that a user should not be allowed to avoid paying for access to a collection that has been developed at great cost and effort by the proprietor. We submit that this concern is likely to be addressed adequately by methods of protection in common use today, including technical restrictions on access and contract. Further, we are concerned that a prohibition directed solely at extraction moves dangerously close to the creation of an intellectual property right in information qua information. As the Supreme Court said in *Feist*, facts are in the public domain; they "may be copied at will;" *it is the means by which the progress of science and art are advanced.* *Feist*, 499 U.S. at 348, 350 (emphasis added).

Notwithstanding this compelling authority, in the interest of compromise, we are willing to work with the Subcommittee to develop an appropriately tailored prohibition on unauthorized extraction when that extraction will cause a substantial injury to the incentive necessary to undertake the investment and effort of creation. We believe the strongest argument for such a prohibition lies within the primary market for the collection. If the primary market is destroyed, there will be little incentive for creation. *See Restatement (Third) of Unfair Competition*, section 38, comment c. ("Appeals to the misappropriation doctrine are almost always rejected when the appropriation does not intrude upon the plaintiff's primary market.") It is not clear that the same is true of subsidiary markets. *See* point III.D., below. Again, we stress that if Congress is to err in setting the scope of this new legislation, it should err on the side of caution and the public domain.

Further, the extraction should not be merely an extraction of some information from the collection, it should be an extraction that captures the sweat and effort of the originator. We discuss this substantiality standard in part III.C., below.

We do not propose that the analysis end with the definition of the scope of prohibited conduct. Just as the Copyright Act contains exceptions that, in appropriate circumstances, permit conduct that falls within one of the exclusive rights, so too should appropriate exceptions be included in legislation to protect collections of information. Thus, there will be highly beneficial non-profit activities and

transformative activities that deserve encouragement and protection. We discuss these exceptions below. However, it is critical that the bill start from a clearer, and narrower concept of prohibited conduct.

## **B. The Legislation Should Clearly Define Protected "Collections."**

The bill protects "collections of information," which are broadly defined as "information that has been collected and organized for the purpose of bringing discrete items of information together in one place or through one source so that users may access them." This definition is so broad that it literally covers almost any publication, including an article, a textbook, the report of a scientific study (with accompanying data), or even, for those of you who are lawyers, a casebook. Each of the foregoing collect information, arguably discrete items of information, for the purpose of making the information accessible. Thus, this definition is inconsistent with our second standard - clarity of subject matter.

Last year's House Report on identical language in H.R. 2652 made clear that the bill is not intended to add to the copyright protection available for articles, texts and reports. The Supreme Court recognized almost 120 years ago in *Baker v. Selden* that

The very object of publishing a book on science or the useful arts is to communicate to the world the useful knowledge which it contains. But this object would be frustrated if the knowledge could not be used without incurring the guilt of piracy of the book.

101 U.S. 99, 103 (1880). It reiterated this identical statement in *Feist*, 499 U.S. at 350. Rather, it is our understanding that the bill is intended to fill a hole in the protection available to large collections that require great effort undertaken for the purpose of gathering and presenting the collection as a collection. Unfortunately, the text of the bill fails to achieve the desired goal. We are not comfortable leaving such an important issue to legislative history.

***Proposed Alternative.*** The definition of protected collections should be modified to make clear that protected collections encompass only compilations that comprise a large number of discrete facts or other items of information collected from numerous sources with the expenditure of substantial monetary or other resources. The term "database" used in the Fair Database Competition Act and Senator Hatch's Discussion Draft more accurately captures the type of protected work than the term "collection of information" used in H.R. 354. Further, the definition should expressly exclude textbooks, articles, biographies, histories, other works of narrative prose, specifications, and other works that include items of information "combined and ordered in a logical progression or other meaningful way in order to tell a story, communicate a message, represent something or achieve a result," as set forth in the House Report on H.R. 2652.

## **C. To Be Actionable, an Extraction Should Take Material that Is Quantitatively and Qualitatively Substantial and Should Appropriate the Effort and Investment of the Creator of the Original Collection.**

The quantum of misappropriation that triggers liability under H.R. 354 is unreasonably broad in two key respects. First, the bill improperly prohibits a taking even if the taking is quantitatively or qualitatively insubstantial. Second, the bill may be read to prohibit a taking even if the portion taken was not the subject of substantial investment (as long as other material in the collection was the subject of such investment). Only a taking that is both quantitatively *and* qualitatively substantial, *and* that appropriates content embodying substantial investment by the creator, should fall within the scope of protection. Any other rule violates the core principle that should govern H.R. 354 - it grants to the originator of a collection the ability to claim a monopoly on facts themselves.

Section 1402 prohibits an extraction or use of a "substantial part" of a protected collection. Pursuant to section 1402, the substantiality standard is satisfied if the part is either quantitatively or qualitatively substantial. This standard grants inappropriately broad protection.

There is no justification for protecting a part of a collection that is not quantitatively substantial. The essence of "collection" that is protected by H.R. 354 is the substantial investment necessary to gather *large quantities* of data or other information from disparate sources. It simply makes no sense to subject a user to potential liability for removing small quantities of material, regardless of how important "*qualitatively*" they may be. Indeed, such small quantities are precisely the individual items of data or information that should not be the subject of protection. *A single fact, two facts or three facts remain simply facts, regardless of how "qualitatively" important they are.*<sup>4</sup>

Likewise, there is no justification for prohibiting the appropriation of even a large quantity of material that is qualitatively insubstantial. The oft-stated rule of *de minimis non curat lex* is as simple as it is wise. The law should not concern itself with trifles. If the extracted material is qualitatively unimportant, there should be no cause of action.

The bill also lacks appears to lack a requirement that there be a nexus between the protected investment in the collection and the taking. It is the investment that should be protected, not the facts themselves. A cause of action should exist only if the defendant has taken content embodying a substantial amount of protected effort. This is best illustrated by example. Consider a large database of information, half of which was taken by the compiler from a pre-existing, public domain database that required very little effort by the creator of the large database. A user whose only extraction is of the pre-existing material should not be subject to liability even if the extracted material qualifies as qualitative and quantitatively "substantial." Granting protection to that which was not the subject of the creator's substantial investment does not serve the purpose this bill seeks to serve.

Moreover, extending protection to such material would extend protection far beyond the analogous protection extended by copyright law. Copyright law protects only the creator's own expressive contribution. Thus, for example, a copyright in a derivative work does not extend added protection to the underlying material on which the derivative work was based. 17 U.S.C. section 103(b). If the underlying material is in the public domain, it is fully lawful to extract the underlying material from the derivative work. Similarly, the protection for a work (including a work that is not derivative work) does not extend to material not created by the author. *See, e.g., Computer Assocs. Intern., Inc. v. Altai, Inc.*, 982 F.2d 693, 710 (2d Cir. 1992) (excluding from copyright protection elements of a work taken from the public domain, each of which "is free for the taking and cannot be appropriated by a single author even though it is included in a copyrighted work"); *Gates Rubber Co. v. Bando Chem. Indus., Inc.*, 9 F.3d 823, 837 (10th Cir. 1993) ("a court must filter out [from copyright infringement analysis] all unoriginal items of a [computer] program, including those elements found in the public domain.").

**Proposed Alternative** . Only a taking that is both quantitatively **and** qualitatively substantial should be prohibited. Further, a cause of action should exist only if the defendant has taken content embodying a substantial amount of protected effort.

#### **D. The Market Harm Standard Is Undefined and Far Exceeds the Injury Cognizable in Unfair Competition Law.**

It is our understanding that the "market harm" standard is intended to be very broad, including as market harm even non-payment of a license fee. Such a standard essentially places the definition of "actual markets" within the control of the proprietor and means that each activity involving information taken from a collection may be unlawful. This standard is inconsistent with long-standing principles of unfair competition and misappropriation law. Instead, it approaches the very concepts of intellectual property law from which this legislation is intended to depart.

If "market harm" means lost license fees, the proprietor merely needs to identify all possible uses and structure a set of licenses to capture different fees for different uses. Even if the proprietor does not at first have such a licensing structure, once a use is discovered or otherwise identified, the proprietor can easily establish a new form of license. At that time, the market becomes an "actual market."

The bill's inclusion of "potential markets" merely exacerbates the problem. The proprietor need not even create the license structure, it need only demonstrate that it might do so.

The concepts of "actual" and "potential" markets are both unacceptably broad and extend far beyond traditional unfair competition and misappropriation law. As the Third Restatement of Unfair Competition states:

the recognition of exclusive rights in intangible trade values can impede access to valuable information and restrain competition. . . . The recognition of exclusive rights may thus deny to the public the full benefits of valuable ideas and innovations by limiting their distribution and exploitation. *In addition, the principle of unjust enrichment does not demand restitution for every gain derived from the efforts of others.* . . . The better approach, and the one most likely to achieve an appropriate balance between the competing interests [between protection and access], does not recognize a residual common law tort of misappropriation."

Restatement (Third) of Unfair Competition, section 38, comment b at 409, 411 (emphasis added). Accordingly, courts

have recognized that broad application of the unjust enrichment rationale in a competitive marketplace would unreasonably restrain competition and *undermine the public interest in access to valuable information.* . . . In most of the small number of cases in which the misappropriation doctrine has been determinative, *the defendant's appropriation, like that in [AP v. INS], resulted in direct competition in the plaintiff's primary market.* . . . Appeals to the misappropriation doctrine are almost always rejected *when the appropriation does not intrude upon the plaintiff's primary market.*

*Id.* at 412-13. In short, traditional misappropriation law remedies harm to plaintiff's "primary market," not "actual" or "potential" markets.

The bill's use of broad market harm concepts departs radically from the misappropriation doctrines that the bill is supposed to adopt. Rather, the use of such broad terms effectively creates a new species of quasi-intellectual property law, one that threatens to access to information.

***Proposed Alternative.*** Concerns over the breadth of market definition can be significantly ameliorated if the actionable conduct proscribed by the bill is limited as described in Part III.A. The bill should prohibit the further dissemination of substantial quantities of extracted material in a manner that is likely to serve as a market substitute for the original in plaintiff's primary market. Further, if extraction of substantial quantities of material is itself to be prohibited, it should be limited to extractions that cause substantial harm to plaintiff's primary market. Substantial harm should not simply mean lost licensing fees.

#### **E. The Exception for Non-Profit Educational Activities Contains a Broad, Vague Condition that Vitiates its Protection.**

Subsection 1403(a) includes an exception for non-profit educational, scientific or research uses that addresses some of the concerns of the education community. However, the exception is limited to those uses that do "not harm directly the actual market for the product or service." Unfortunately, as discussed in Part III.D., above, this limitation may be so broad that it destroys the exception of much of its value.

As discussed in Part III.D., the term "actual market" is infinitely flexible, and can mean anything the proprietor wants it to mean. All the proprietor must do is declare the existence of a license to cover a particular use or type of use, and it has created an "actual market." Further, there is no explanation of how "direct harm" differs from any other kind of harm. The House Report on H.R. 2652 stated that even loss of license fees was intended to constitute a "direct" harm.

Further, because the provision is crafted as an exception, the burden of proving that conduct qualifies for protection likely falls upon the educational institution. The burden should fall on plaintiff to demonstrate that wrongful conduct has occurred.

In other words, subsection 1403(a) provides a database proprietor with sufficient ammunition to eliminate the exception entirely, and foreclose the very educational, scientific and research activities that should be preserved. At the very least, the application of the exception to any case of nonprofit use can be subjected to costly litigation, which itself threatens to chill these important activities.

***Proposed Alternative.*** To ensure that nonprofit educational, scientific and research activities receive appropriate protection, two changes should be made to the bill. First, the scope of protection should be narrowed, as discussed in Parts III.A. through III.D., above. This will ensure that educational, scientific and research activities are only subject to liability when they are the primary market for the original collection of information and the activity causes substantial harm to that market. To make this absolutely clear, we would support language such as "under no circumstances shall non-profit educational, scientific or research activity that would otherwise violate this chapter constitute a violation of this chapter if plaintiff fails to bear the burden of proving that the activity is causing substantial harm to the primary market for its protected collection of information." Second, even in these circumstances, courts should have the discretion, through consideration of a flexible "reasonable use" exception, to consider the facts and circumstances of the particular case. *See* Part III.F., below.

#### **F. The Exception for Other Reasonable Uses Is Insufficiently Flexible, and Contains Conditions that Greatly Limit its Benefit.**

The exception in paragraph 1403(a)(2), for "other reasonable uses," represents an improvement over H.R. 2652. However, it is constrained by absolute conditions that restrict its usefulness and that limit its application beyond the considerations applied in traditional copyright fair use analysis. Thus, it does not serve its intended purpose and results in protection for compilations that is greater than that provided by copyright.

Unlike the fair use exception in copyright law, which provides courts flexibility to consider all relevant factors for virtually any kind of use, the "reasonable uses" provision contains absolute conditions that restrict that flexibility. A more flexible approach, coupled with a narrowing of section 1402, is necessary.

***Individual Acts.*** The most notable limitation in paragraph 1403(a)(2) is the limitation to "an individual act of use or extraction of information." An "individual act" is defined as 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." Of course, by its nature, scholarship, teaching, and research require repeated acts and patterns of acts. Thus, many of the acts that should be covered by a reasonable use exception fall outside of the scope of the exception as written.

There is no justification for limiting this exception to a single act of extraction or use. The extraction of a quantum of information has the same impact regardless of whether it occurs in a single act, or as part of several acts. Courts have the flexibility to aggregate multiple acts that are part of a pattern or repeated practice by related parties in their consideration of the exception. That should be sufficient to ensure that the exception is not abused.<sup>5</sup>

***Amount of Extracted Material.*** The exception is not available if the amount extracted is more than is "appropriate and customary" for the purpose. Although this is a factor considered in copyright fair use analysis, it is not an absolute condition for application of the fair use doctrine. It is easy to envision a situation in which more than "appropriate and customary" is extracted, but the excess extraction does not have a significant effect on the market, or may have been extracted by accident or on the basis of a

misapprehension of what was needed. Alternatively, the particular purpose may not have a "customary amount." There is no justification for courts to be required to bar use of the exception if the amount extracted exceeds an amount that is "appropriate and customary." As in copyright fair use, the amount of the taking should be a factor that is considered like all other relevant factors.

***Market Substitution.*** The exception may not be applied if the extracted material is offered in commerce and is likely to serve as a market substitute for the collection. Such activity is not an automatic, absolute bar to a finding of copyright fair use. It should not be an automatic, absolute bar to the reasonable use exception here.

Every day, material is extracted from collections and used in research and other academic work. The extracted material is often transformed by the extensive effort and energy of the original extractor, who develops, selects, and organizes the material to make it useful to others in the field. The resulting work could be a breakthrough in science, social science, philosophy, economics, medicine or any one of scores of other fields. It may not even take the form of a compilation or database. The resulting work may have far greater value to others in the field than the original, undigested database. As a result of the work, others may not need to duplicate the extractor's efforts, and the resulting work may "substitute" for the broad, undigested database. This is precisely the type of work that should be encouraged, not sanctioned.

***Proposed Alternative.*** A "reasonable extraction and use" exception should be crafted which is similar to the copyright fair use exception and which does not contain absolute bars to its application. For example, the factors should include the nature of the use, amount of the material extracted or used, and effect on the market for the work. Sub-factors may be identified, including (i) the extent to which the extraction or dissemination is commercial or nonprofit, (ii) if the extracted material is incorporated into an independent work or collection, the extent of transformation, (iii) whether the amount of material extracted is appropriate and customary for the purpose of the extraction, (iv) whether any resulting collection is marketed to persons engaged in the same field or business as the extractor, and (v) the extent to which a resulting collection that is offered in commerce is likely to serve as a significant market substitute in the primary market for the original collection. Courts should also be instructed to take into account that the nature of the collection, as a collection of facts, is ordinarily entitled to less protection against fair use than an original work of creative authorship under copyright law.

This reasonable use exception should be combined with the narrower scope of protection discussed in Parts III.A. through III.D.

#### **G. The Bill Should Protect the Public against Unreasonable Market Power from Compilations that Are Not Readily Available from Competitive Sources.**

The public ordinarily is protected from unreasonable prices arising from a seller's unconstrained market power through the operation of the marketplace. If excessive prices are charged, large profits will be earned, which in turn will encourage the entry of competitors. Unfortunately, there are many reasons that this mechanism might not work for compilations of information. The information contained in a compilation may not be readily available from competing sources. It may be under the control of the compilation provider; it may be historical information that is no longer available to the public; or it have been collected at substantial cost over a long period of time, creating a barrier to competitive entry. In these cases, providing the protection contemplated by H.R. 354 could grant substantial market power, resulting in excessive prices for information and returns greater than that needed to stimulate development of the information products. As the Federal Trade Commission cautioned in its September 28, 1998 comments on last year's Collections of Information Antipiracy Act, "policies the further entrench the market power of single-source data providers could have an unintended potential for anticompetitive conduct."

In its consideration of granting these new rights over information products, Congress should ensure that such power is not created. There is no public policy justification for granting super-competitive

market power to the providers of information products. The incentive to create will be maintained by providing reasonable returns, not monopoly returns. Conversely, as discussed at the beginning of this statement, there are strong public policy reasons to ensure that information is made available to the public on a reasonable basis.

Antitrust law does not provide adequate protection against market power that may be created by legislation. The Federal Trade Commission recognized this point in its comments last year. Market power is not per se unlawful under the antitrust laws. It may be gained in numerous ways that are lawful, including the granting of rights by Congress. Even if charging high prices or imposing unreasonable terms did violate the law, it is extremely difficult and expensive to bring an antitrust case. Such suits are not viable protection from the market power that this bill creates.

***Proposed Alternative.*** There are several possible approaches to address the issue of unreasonable market power. For example, the bill could encourage database proprietors to charge reasonable royalties by providing that it shall be a complete defense to an action that plaintiff did not make the compilation available for a reasonable royalty. Alternatively, the bill could limit the available remedy to a reasonable royalty in cases in which competitive sources of the information in the compilation did not exist. These limitations should apply with respect to any extraction right that is included in the bill and to the incorporation of information into a transformed product (to the extent such conduct remains actionable under section 1402).

#### **H. The Bill Should Secure Access to Older Versions of Protected Compilations, in order To Protect Against Perpetual Protection.**

The revision of section 1408(c), which provides that investment in the maintenance of an existing compilation cannot extend the term of protection for that compilation, is a valuable addition to the bill. We are concerned, however, that it is only a partial solution to ensuring that the bill does not result in perpetual protection for continually updated compilations. It does little good for the bill to end protection for an old compilation if the compilation is no longer accessible by the public. Access to older versions of compilations that have fallen into the public domain should be preserved.

***Proposed Alternative.*** One means of preserving access, of course, is to create a registration and deposit system using the Copyright Office. This approach would have the advantage of the Copyright Office's experience in protecting access and would ensure that protection would be contingent on access following expiration of protection. However, this approach would entail administrative costs that Congress may not wish the public to bear.

Another approach that could work would be to require the party seeking protection under the Act to maintain expired versions of the compilation in a manner freely accessible to the public. Failure to comply with this requirement would preclude protection for more recent revised versions of the compilation. In order to ensure that the public is aware that earlier versions are available, the proprietor of a database should be required to include a notice with the compilation setting forth the year protection was first claimed and, after older versions become available, where they may be found.

#### **I. The Bill Should Ensure that Institutions that Act as Online Service Providers Are Not Subjected to Liability.**

As the higher education community pointed out in last years' debate on H.R. 2281, the Digital Millennium Copyright Act, universities and colleges frequently offer their students and faculty access to the Internet and other computer networks. Congress recognized that these institutions could not reasonably be held accountable for the online conduct of such persons who infringe copyrights. It is even more difficult for an institution providing such online services to differentiate between the facts and information that form the basis of academic discourse, and protected "compilations."

Unfortunately, as it is now drafted, a system that transmits or stores a collection or a part of a

collection could be said to be "using" that collection. Even narrowed as recommended in Part III.A., a system that is used to disseminate extracted material could be said to be participating in the dissemination. Such a result would not be appropriate, and for all of the reasons discussed in the debate last year, should be expressly precluded.

The prohibitions of H.R. 354 should be targeted at the person or entity engaging in the specifically proscribed conduct, not on intermediaries whose systems or networks are used by those persons or entities to transmit information. The bill should make clear that no liability attaches to such intermediaries.

***Proposed Alternative.*** The Senator Hatch's Discussion Draft and the Fair Database Competition Act both include a provision making clear that online service providers are not subject to liability. This provision should be included in H.R. 354. In addition, universities and colleges need language making clear that the conduct and knowledge of faculty members and graduate students engaged in research or teaching activities will not be imputed to the institution.

#### **J. Non-Profit Teaching Activities Should Be Granted a Clear Exemption from Liability, Akin to the Exemption in Copyright Act Section 110.**

Non-profit teaching activities are entitled to clear exemptions from liability under Copyright Act section 110(a)(1) & (2). These provisions of copyright law make clear that it is not an infringement to perform, display or transmit a copyrighted work in the course of instruction by a non-profit educational institution. These exemptions are not subject to the risk of fact-intensive litigation that surrounds more general claims of fair use. In keeping with the principle that H.R. 354 should not offer protection that is not available to copyrighted works under copyright law, similar exceptions should be provided for compilations protected by H.R. 354.

***Proposed Alternative.*** The extraction and dissemination of information for purposes of display or distribution to pupils in the classroom (or the distance education equivalent) should be granted a clear exception from liability.

#### **IV. Technical Concerns**

In addition to the major substantive concerns discussed above, the Higher Education Associations have a few comments of a more technical drafting nature. We assume that the first two are simply drafting oversights.

##### **A. Overly Broad Monetary Relief.**

The bill provides for monetary relief that is significantly greater than that provided under copyright law. As drafted, subsection 1406(d) mandates recovery of "defendant's profits not taken into account in computing the damages sustained by the plaintiff." Curiously, however, the subsection does not contain a critical limitation contained in Copyright Act section 504(b), after which it was modeled. Specifically, recoverable profits are limited to those "that are attributable to the infringement." Similarly, subsection 504(d) permits a defendant to prove and deduct "the elements of profit attributable to factors other than the copyrighted work." The analogous language is not found in subsection 1406(d). It is inconceivable that H.R. 354 contemplates the recovery of *all* of the violator's profits, relief that far exceeds that available under copyright law.

##### **B. Limitations on Penalties and Criminal Liability for Non-Profit Educational Institutions.**

The Higher Education Associations appreciate the two limitations of liability for employees and agents of non-profit educational institutions contained in the bill. Section 1406(e) provides for the reduction of monetary relief in cases in which the defendant is the employee or agent of a non-profit educational institution acting in good faith. Section 1407(a)(2) precludes criminal penalties against the

employee or agent of a non-profit educational institution. In each case, the provision applies to employees or agents of educational institutions, but inexplicably fails to provide equivalent protection to the institution itself. It is certainly possible that suit will be brought directly against a non-profit educational institution in its own right, rather than against an employee or agent. Each of these limitations should expressly apply to the institution as well as to its agents and employees.

### **C. The Protection of State University Databases under Section 1404(a).**

As we noted at the outset of this statement, universities and colleges are not only users of compilations of information, they also act as creators of collections that should be protection to the same extent as collections created by commercial providers. In this statement, we have argued for specifically targeted protections for collections of information. Whatever level of protection Congress ultimately deems appropriate should be available to universities and colleges on the same terms. Moreover, there is no reason to discriminate against state universities and colleges when they act as creators. For these reasons, we appreciate the recognition that collections developed by state educational institutions are not precluded from protection under section 1404. We are concerned, however, that the language "in the course of engaging in education or scholarship" might be misconstrued to exclude research or other activities of the institution. For that reason, we suggest that the final nine words of paragraph 1404(a)(1) be stricken.

---

<sup>1</sup> The term "use" would also appear to cover activities, such as the internal creation of derivative compilations, that might be within the scope of copyright rights when applied to original works of authorship, but which should not be within the protection afforded to collections of information. Nor is "use in commerce" likely to be a meaningful limitation, given the breadth of the term "commerce." Any entity whose use of information from a collection of information involves an instrumentality of interstate or foreign commerce (e.g., telephone, mails, or the Internet) would be acting "in commerce."

<sup>2</sup> We discuss the types of "markets" that should be considered in Part III.D., below.

<sup>3</sup> We understand that there are objections that the competition bill's scope is too narrow, and are prepared to work with the Subcommittee to satisfy its concerns.

<sup>4</sup> Subsection 1403(b) also reflects H.R. 354's problems in this regard. Although this subsection appears intended to foster the principle that individual facts are not protected, it may actually have the opposite effect. The first sentence states that the bill does not "prevent the extraction or use of an individual item of information or other insubstantial part." The second sentence provides that an individual item of information shall not, itself, be considered substantial. Of course, stating that insubstantial parts of a collection are not prohibited adds nothing to section 1402 which purports to prohibit only the taking of a "substantial part" (however broadly that is defined). We are concerned that by explicitly stating that a single item of information may never be substantial, subsection 1403(b) implies that two or three items of information may be substantial. A better approach would be to tighten the substantiality standard of section 1402 and make clear that the extraction of a relatively small portion of the collection would not be deemed substantial.

<sup>5</sup> At most, the exception could provide that "acts that are 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 should be aggregated for consideration of the applicability of this paragraph."

---

- [ARL Federal Relations and Information Policy](#)
- [Copyright and Intellectual Property Table of Contents](#)

---

[E-News](#) [Recent](#) [Copyright](#) [Govt. Info](#) [Telecom.](#) [OtherIssues](#) [GIS](#) [Letters](#)



© Association of Research Libraries, Washington, DC

Web Design by [Michael Miller](#)

Maintained by [ARL Web Administrator](#)

**Last Modified:** March 19, 1999