

Statement of
ANDREW J. PINCUS
GENERAL COUNSEL
UNITED STATES DEPARTMENT OF COMMERCE
before the
SUBCOMMITTEE ON COURTS AND INTELLECTUAL PROPERTY
COMMITTEE ON THE JUDICIARY
U.S. HOUSE OF REPRESENTATIVES
concerning
H.R. 354, the "COLLECTIONS OF INFORMATION ANTIPIRACY ACT"
18 March 1999

[I. Introduction](#)
[II. History of Administration Study of Database Issues](#)
[III. General Principles](#)
IV. Discussion

- [A. The First Principle](#)
- [B. The Second Principle](#)
- [C. The Third Principle](#)
- [D. The Fourth Principle](#)
- [E. The Fifth Principle](#)
- [F. The Sixth Principle](#)
- [Additional Issues](#)

[Appendix A](#)
[Appendix B](#)
[Appendix C](#)
[Appendix D](#)
[Notes](#)

Mr. Chairman and Members of the Committee:

Thank you for this opportunity to present the Administration's views on H.R. 354, the "Collections of Information Antipiracy Act."

I. INTRODUCTION

The Administration views database protection legislation from a number of perspectives: as a creator of data and a user of it; as an advocate of both economic incentives for socially useful investment and open, market-based competition free from artificial barriers; and as an entity committed both to effective law enforcement and to the First Amendment. Reconciling these perspectives is difficult in any context. The digital economy's rapid and unpredictable change makes this challenge even greater.

The Administration believes strongly in free markets, in which firms can meet demand for new products and services without having to overcome artificial barriers that keep consumers hostage to an undesirable status quo. However, we also recognize that there are circumstances in which markets need legal mechanisms in order to function efficiently. The *Feist* decision⁽¹⁾ conclusively eliminated one form of legal protection for databases. Undeniably, *Feist* has altered the landscape, but the topography is still changing in ways that pull in different directions as to the nature and extent of protection that is needed.

In particular, the emerging digital environment has significant implications for this issue. It has become commonplace to observe that information is the currency of our economic age. That puts a premium on designing a legal schema that creates sufficient incentives to maximize investment in data collection -- to expand the available universe of information -- without putting in place unjustified obstacles to competition and innovation. Moreover, digital technology permits the creation and distribution of a large number of perfect copies of data files at the touch of a button and therefore expands dramatically the risk that, in the absence of adequate legal remedies, piracy, or the threat of piracy, will deter investment in database creation. For all of these reasons, it is important to calibrate new private rights carefully -- to optimize overall economic and social benefits, to prevent unfairly undermining investments and agreements premised on the current law, and to preclude new opportunities for thwarting competition.

The U.S. Government has a unique stake in database legislation because it collects, manages, and disseminates massive amounts of information, possibly more information than any other entity in the world. In all these processes, it interacts with the private sector in a variety of ways. In addition, federal agencies are engaged in funding research that produces tremendous amounts of information that the government does not undertake to manage itself.

These activities represent enormous investments in highly complex knowledge management processes that are vital to human health, the environment, national security, scientific progress, and technological innovation -- and, in turn, to the economy as a whole. Changes in ground rules for the use and reuse of information must be designed to minimize disruption of these critical activities and to avoid imposition of new costs that could hinder research.

The sections which follow discuss the Administration's efforts to study database protection and access issues (Part II) and summarize the six principles that we believe should guide both domestic legislative and international treaty efforts in this area (Part III). Next, we elaborate on each principle, discussing the Administration's concerns relating to that topic and the range of possible solutions on which we believe interested parties should focus (Part IV). Finally, we offer some additional points that should be included in any database protection legislation.

II. HISTORY OF ADMINISTRATION STUDY OF DATABASE ISSUES

In response to legislative proposals in Congress and developments in the World Intellectual Property Organization (WIPO), the Administration devoted substantial energy in 1998 and 1999 to studying database protection and access issues. The Administration's review of these issues has included a variety of mechanisms and fora:

- The Patent and Trademark Office (PTO) held a public conference on database protection and access issues on April 28, 1998.

- During the spring and summer of 1998, a variety of Executive Branch departments and agencies participated in an informal working group on database issues led by the State Department, the Office of Science and Technology Policy (OSTP), and the PTO.
- In January 1999, the National Research Council held a two day conference on scientific databases at the Department of Commerce. This conference was supported by the National Science Foundation, the National Institutes for Health, and several other agencies.[\(2\)](#)
- Various officials in the Executive Office of the President (including OSTP), the Department of Commerce (including PTO), and the Justice Department have held informational meetings with both proponents and opponents of database protection legislation.

In addition to these efforts, the Administration has carefully studied a wide range of reports, studies, legal opinions and legislation on database protection and access from the United States, Canada, Japan, and the European Union, as well as participating in discussions of database protection issues at WIPO conferences in 1996, 1997, and 1998.

The Administration continues to discuss these issues with concerned parties and to examine specific topics and areas where we believe further information will help both the legislative process and any future study of the effects of database protection that might be mandated by legislation.

III. GENERAL PRINCIPLES

On August 4, 1998, in response to Senate consideration of then-H.R. 2652, the Administration set out the principles that it believes should govern database protection legislation.

Now, as then, Administration supports legal protection against commercial misappropriation of collections of information. We believe that there should be effective legal remedies against "free-riders" who take databases gathered by others at considerable expense and reintroduce them into commerce as their own. This situation has arisen in recent case law, and we believe that digital technology increases opportunities for such abuses.

At the same time, the Administration's concerns with the provisions of H.R. 354 are similar to those we expressed with respect to H.R. 2652, including the concern that the Constitution imposes significant constraints upon Congress's power to enact legislation of this sort. From a policy perspective, the Administration believes that legislation addressing collections of information should be crafted with the following principles in mind:

1. A change in the law is desirable to protect commercial database developers from commercial misappropriation of their database products where other legal protections and remedies are inadequate.
2. Because any database misappropriation regime will have effects on electronic commerce, any such law should be predictable, simple, minimal, transparent, and based on rough consensus in keeping with the principles expressed in the Framework for Global Electronic Commerce. Definitions and standards of behavior should be reasonably clear to data producers and users prior to the development of a substantial body of case law.
3. Consistent with Administration policies expressed in relevant Office of Management and Budget circulars and federal regulations, databases generated with Government funding generally should not be placed under exclusive control, *de jure* or *de facto*, of private parties.
4. Any database misappropriation regime must carefully define and describe the protected interests and prohibited activities, so as to avoid unintended consequences; legislation should not affect

established contractual relationships and should apply only prospectively and with reasonable notice.

5. Any database misappropriation regime should provide exceptions analogous to "fair use" principles of copyright law; in particular, any effects on non-commercial research should be *de minimis*.
6. Consistent with the goals of the World Trade Organization (WTO) and U.S. trade policy, legislation should aim to ensure that U.S. companies enjoy available protection for their database products in other countries on the same terms as enjoyed by nationals of those countries.

With these principles in mind, we turn to an analysis of H.R. 354.

IV. DISCUSSION

A. First Principle -- Protect against commercial misappropriation

A change in the law is desirable to protect commercial database developers from commercial misappropriation of their database products where other legal protections and remedies are inadequate.

The Administration supports enactment of a statute to protect database creators against free-riding --- the wrongful taking and distribution of database material with resulting infliction of commercial harm (loss of customers) on the database creator. Indeed, there is considerable, if not complete, consensus that this kind of free-riding can occur without additional legal protection for non-copyrighable databases and that such legal protection is necessary to prevent a diminution in database creation.⁽³⁾

Section 1402 is the operative core of H.R. 354, providing the "basic prohibition" of this proposal to protect collections of information through a misappropriation model.⁽⁴⁾ Section 1402 prohibits unauthorized commercial misappropriation of a substantial amount of a database; it also appears to prohibit unauthorized extraction or "use" of data from a database by an individual, no matter how the information is used.

We do not believe that protection of that breadth is appropriate in the database context. As a policy matter, we must weigh the need to protect database creators against the potential impact on scientific research in particular, and the dissemination of information within the society generally. It therefore makes sense to focus any prohibition on the precise activities that pose the commercial threat -- "use" is simply too broad and ambiguous.⁽⁵⁾ Indeed, the breadth and ambiguity of the prohibition has required concerned parties to focus considerable attention on expanding the list of statutory exceptions to make clear that various activities would not be affected by the prohibition. We believe it more appropriate to narrow the prohibition so it is targeted on conduct like the troubling acts of commercial misappropriation identified in the *Warren Publishing* and similar cases.⁽⁶⁾

H.R. 354's basic prohibition consists of three basic elements, imposing liability on any person who "extracts or uses in commerce" all or a substantial part of a database so as to cause "harm" to the "actual or potential market" of the database creator. In our view, all three of these elements should be focused more precisely on the commercial free-riding situation.

To begin with, the "extract[s] or use[s]" language should be narrowed. One approach would be to limit the reach to a person who, without authorization "extracts for commercial distribution or distributes in commerce" all or a substantial part of a database. The substitution of "distribution" in place of "use" would clarify that the Act is directed at active behavior, rather than receptive activities such as viewing, reading, or analyzing. "Extracts for commercial distribution" would cover any replication preparatory to

distribution in commerce. Distributes in commerce should be understood broadly, compatible with First Amendment concerns.

While the Administration continues to believe that misappropriation for commercial purposes should be the focus of any legislative efforts, we recognize that, when systematic, some acts that might be characterized as "extraction" (in other words, acts of duplication) by individuals could conceivably undermine the commercial market for a database product. We are not familiar with any reported cases or incidents of this kind, but we recognize that such harm could occur. Such damage may occur when those acts becomes customary in a particular economic sector or field of research. At present, if there is no contract with the individual or his/her organization, the investor in a database has no effective civil remedy against such acts.⁽⁷⁾ We believe that one of the greatest challenges in drafting database protection legislation is providing database producers with some type of protection against such *patterns* of repeated individual activity without prohibiting uses of data by individuals that most people believe should be treated as "fair uses" and without violating the First Amendment. We are not certain whether a balance can be struck. Our suggested language concerning "extraction for distribution" and "distribution" does not address this issue; we look forward to working with the Subcommittee on this matter as the legislation moves forward.

Second, the Subcommittee should consider whether the requirement of "harm" in section 1402 should be elevated to "substantial harm" as a means of shielding *de minimis* activities from any possible liability. We know that some proponents of H.R. 354 have expressed concern about a "substantial harm" standard because they believe that judges would compare the standard unfavorably to copyright law, which requires only "harm." We agree that it is important to anticipate how judges would administer any new law, but we believe that a "substantial harm" standard is familiar to courts from other areas of American law.⁽⁸⁾ Appropriate legislative history could direct judges away from unintended comparisons to copyright law or areas of the law where "substantial harm" has been interpreted to impose a higher standard than intended in this bill.

At the same time, some critics of H.R. 354 have suggested that the proper trigger for liability is whether the misappropriation "so reduce[s] the incentive to produce the product or service that its existence or quality would be substantially threatened," a test from the *National Basketball Association v. Motorola* case.⁽⁹⁾ While we agree that a misappropriation law should be focused on acts that do, in fact, have a tendency to reduce incentives in this manner, we think this "diminution of incentive" test is ill-suited as a component of the basic prohibition; it does not comport with the Administration's principle (described below) that a database protection law should be predictable, simple, and transparent. Because a database user cannot be expected to know much about the incentive structures that lead to production of databases, such a user would have no way to judge in advance whether or not her acts would satisfy a "diminution of incentive" test for liability. We also are concerned that the "diminution of incentive" test requires much more complicated proofs than would be incurred with a harm test.⁽¹⁰⁾ Accordingly, we believe that Congress should instead rely upon a "substantial harm" test or similar measure to serve as a workable proxy for the "diminution of incentive" test.

Third, we suggest reexamination of the concepts of "actual" and "potential" market. We are very concerned that, as presently drafted, these concepts are broader than market definitions used in other areas of the law, could be subject to manipulation by private entities, and could too easily expose legitimate business practices to substantial liability. We urge the Subcommittee to consider an objective definition tied to the product's current actual customer base or the market currently exploited by similar products or services. We are concerned that any broader definition might deter entrepreneurs from developing new products and services that add significant value and do not compete directly with the original database. Leaving room for transformative uses is critical in shaping the definition of the market as targeting the free-riding we wish to prohibit. We believe that the Subcommittee should consider, individually and perhaps in combination, the notions of "principal market" drawn from unfair competition law, and "neighboring market" proposed in the Senator Hatch draft.

The Department of Justice notes that this legislation raises serious constitutional concerns that current copyright law does not raise. The Constitution itself provides for protection of copyright, in order to promote progress in science and the arts. Therefore, copyright and the First Amendment are intended to protect analogous values, and are aimed, in part, at similar and compatible objectives. The Copyright Clause and the Copyright Act permit protection only of an author's original expression, and do not authorize protection of facts. This comports with First Amendment principles. By contrast, the proposed prohibition in H.R. 354 would be directed against dissemination of facts. That measure as currently drafted likely would not survive constitutional scrutiny, at least in numerous applications. The constitutional concerns are related to the scope of the basic prohibition, discussed above, as well as the issues discussed below, including the range of permitted uses, resolution of the "perpetual protection" problem and the possibility of "sole source" situations, but the constitutionality of any law in this area will depend upon the particular statutory language adopted and therefore cannot be analyzed definitively at this time. We look forward to working with the Subcommittee to avoid constitutional infirmity.

B. Second Principle -- Keep it simple, transparent, and based on consensus

Because any database misappropriation regime will have effects on electronic commerce, any such law should be predictable, simple, minimal, transparent, and based on rough consensus in keeping with the principles expressed in the Framework for Global Electronic Commerce.⁽¹¹⁾ Definitions and standards of behavior should be reasonably clear to data producers and users prior to the development of a substantial body of case law.

This principle informs all of our analysis. We believe that database legislation should be directed squarely at behavior that is widely acknowledged to be unfair and has been documented as a problem worthy of a legislative response. This will ensure that the legal system is not used to threaten litigation in borderline cases in a manner that may inhibit the flow of factual information and the vigor of free market competition.

We also believe that in introducing this new form of protection, some of the burden of promoting transparency and predictability should be borne by those who benefit. The legislation should not create an environment in which many kinds of database users must suddenly act at their peril. In particular, the Subcommittee might consider how a notice system could effectively warn database users when a database producer is asserting protection under the law. This will also help reduce the costs of identifying multiple cascading interests that are likely to aggregate more frequently in databases than in works of authorship. In this regard, we applaud the addition of the "good faith" of the defendant as a factor in allowing "permitted use" under section 1403 although, for reasons discussed below, we believe that our additional changes to the "permitted use" section may be needed.

Rather than prescribing a particular approach and trying to address the difficulties of implementing it in legislation, we would prefer (again, in keeping with the Framework for Global Electronic Commerce principles) to assign to database publishers and users the responsibility of devising appropriate standards to identify and assert interests. While in paradigmatic cases, such as the circumstances in *Warren Publishing*, there may be no question about deliberate free-riding, the principle remains: users must have reason to believe that their acts are damaging to others. The "good faith" factor in H.R. 354 combined with private sector-developed standards for notice and disclosure would help ensure that the legislation works to condition behavior based on reasonable expectations and to avoid traps for the unwary.

C. Third Principle -- Preserve access to government data

Consistent with Administration policies expressed in relevant Office of Management and Budget circulars and federal regulations, databases generated with Government funding generally should not be placed under exclusive control, *de jure* or *de facto*, of private

parties.

1. Exemption of government data

The U.S. Government collects and creates enormous amounts of information, possibly more than any other entity in the world. State and local governments in the United States also gather and generate tremendous amounts of data. Broadly defined, government-generated data touches every sector of the economy and civic life. Government-funded data ranges from crime statistics to data on subatomic particles; from geological maps to court opinions; from immigration statistics to digital images of distant galaxies.⁽¹²⁾

The Administration believes that a database protection law generally should not protect government investment in generating data. There are three reasons for this conclusion. First, database protection proposals are premised on the need to provide an *incentive* for investment in data gathering; in the case of government-funded information, no incentive is needed. If a government decides that it is in the public interest to collect information on smog levels, education scores, or solar flare activity, it will do so. Second, there is a widespread sentiment that once data generation has been paid for with government funds, taxpayers should not have to pay "twice" for the same data.

Finally, the U.S. Government has historically pursued policies that strongly favor public funding of the creation and collection of information. The Administration believes that these policies have contributed greatly to the success of America's high technology and information industries as well as the strength of our democratic society. The Administration has stated elsewhere:

"Government information is a valuable national resource. It provides the public with knowledge of the government, society, and economy -- past, present, and future. It is a means to ensure the accountability of government, to manage the government's operations, to maintain the healthy performance of the economy, and is itself a commodity in the marketplace."⁽¹³⁾

The Administration believes that the free flow of government-generated data is an important engine of economic growth; it will be an increasingly important resource for any society intent on creating jobs, businesses, and wealth in the "Information Age." Often, government-generated information is also critical to the health and safety of the population; we must ensure that any database protection law does not hamper the dissemination of such information.⁽¹⁴⁾

H.R. 354 addresses the issue of government-generated data with the following section 1404(a) exclusion:

"Protection under this chapter shall not extend to collections of information gathered, organized, or maintained by or for a government entity, whether Federal, State, or local, including any employee or agent of such entity, or any person exclusively licensed by such entity, within the scope of the employment, agency, or license. Nothing in this subsection shall preclude protection under this chapter for information gathered, organized, or maintained by such an agent or licensee that is not within the scope of such agency or license, or by a Federal or State educational institution in the course of engaging in education or scholarship."

The Administration believe that this provision serves the general policy goal of making all forms of government information available to the public, but we believe the language is too narrow to satisfy this goal fully.

To begin with, we suggest that the Subcommittee examine existing definitions of "government information" for more inclusive descriptions of government-sponsored data collection. For example,

OMB Circular A-130 states that "the definition of 'government information' includes information created, collected, processed, disseminated, or disposed of both by and for the Federal Government."⁽¹⁵⁾ In particular, we believe that the present language does not adequately cover situations in which the government contracts for or provides grants for information gathering. For example, for reasons of accountability, several government contracts expressly state that the private entity is not an "agent" or "licensee" of the government, removing the data gathering from the ambit of section 1404(a). One way to address this would be to include language that information collected "under government contract, grant, or other agreement" is covered by section 1404(a). Another possibility would be inclusion of language making clear that the 1404(a) exclusion also applies to data gathering "funded by the government."

In crafting broad statutory language that includes works created by government contract as government collections of information, a distinction should be drawn between (a) compilations of data made as a necessary element of a government-funded activity, and (b) compilations of data made by private entities over and above the activity being funded by the government. This appears to be the intent of the section 1404(a) language that:

"Nothing in this subsection shall preclude protection under this chapter for information gathered, organized, or maintained by [a government] agent or licensee that is not within the scope of such agency or license . . ."

This test also should be modified to account for government contractors and grantees who are neither licensees nor agents. In addition, standards for when preparation of a database is mandated by government contract could be developed from existing standards for when government agencies must collect data.⁽¹⁶⁾

We also note that 1404(a) is currently worded so that data gathered by state-funded colleges and universities may enjoy protection under the bill. This same provision appeared in H.R. 2652 and the Committee report for that bill indicated that the statutory language was intended to ensure that "institutions that happen to be government owned should not be disadvantaged relative to private institutions when producing databases . . ." The Administration respectfully disagrees with this reasoning; we believe that public universities should fall within a broad definition of government institutions which generate collections of information. Instead of trying to draw a distinction between public universities and other government institutions, it might be more appropriate to concentrate on the distinction between *public* research and *privately funded* research at *public* institutions.⁽¹⁷⁾

Higher education institutions are also a fertile ground for situations in which a database's generation is *partially* funded by the government. In such circumstances, what is fair to the researcher and to the public? The Senator Hatch discussion draft would have placed outside the protection regime those databases "the creation or maintenance of which is substantially funded by [a] government entity."⁽¹⁸⁾ Without conducting a detailed analysis of the Senate discussion draft provisions, we believe in general that databases produced with substantial government funding should be treated like databases of government-generated data, at least in the absence of a specific contrary provision in the government contract, grant, or other agreement.

2. Dissemination of government-generated data and the potential for "capture"

Once data has been generated with public funding, there remains the goal of disseminating that data as broadly as possible. For many government agencies, the responsibility to make government-generated information widely available is a statutory obligation.⁽¹⁹⁾ Dissemination of government-generated data has always involved a mix of public and private resources. Through the Congressionally mandated Federal Depository Library Program, the Federal Government uses public libraries, libraries of public universities, and libraries of private institutions to make government-funded

information widely available to citizens. In hundreds of cases ranging from the court system to the U.S. Geological Survey, private entities gather raw, government-generated data and then process, verify, and repackage the data to produce value-added products which are then widely disseminated.

Once there are such commercial products, any decisions to devote public resources to disseminate the raw government data further must be weighed against other demands for government resources.⁽²⁰⁾ If government-generated data does not remain available to the public from government sources, there is the potential for capture of data, with one or a few private entities becoming the "sole source" for important data.

When a U.S. Government work is integrated into a private, value-added product, copyright law requires that the U.S. Government portion remain unprotected and available for copying.⁽²¹⁾ The Administration has considered whether a parallel solution to the "capture" problem with collections of information would be appropriate: requiring private entities to identify government information in their value-added products, and excluding such information from any database protection schema. The problem with this approach is that a private entity may make a considerable investment in gathering government data from disparate sources, bringing it together, and distributing it. This "value-added" would be lost - and the incentive for it destroyed - if all the data could be freely appropriated on the grounds that it is government-generated data in a private database.

On the other hand, not requiring that the government-generated data integrated into a private product remain outside the database protection schema creates the risk of "capture." Many people believe that this is a significant danger in the case of published court opinions in which there are only two major private publishers.⁽²²⁾ Even when government-generated data remains available to the public from the government, it may be much more difficult to obtain than the private, value-added product. If only because the government does not advertise, it may *appear* that the private entity is the sole source for the government-generated data (both in the raw or value-added form).

The Administration does not have any single proposal that will solve all of these issues. We do, however, have a few specific suggestions to address, to some degree, the capture and sole-source problems with government-generated data.

First, we recognize the importance of keeping government-generated information in the public domain, and urge agencies whose grants, contracts, or other agreements involve a significant amount of data generation to include provisions in the grants, contracts, or other agreements that require grantees, contractors, and the like to make research results available to the public in a non-commercial form. The Administration would support language calling for a study to address this issue and offer recommendations to agencies, either individually or collectively, on how to improve non-commercial access to government-generated data resulting from research. At the same time, our recent experience with legislative mandates to amend OMB Circular A-110 counsels against any attempts at this time to impose any uniform access requirements on the wide range of government agencies.⁽²³⁾

Second, we believe that any database protection law along the lines of H.R. 354 should require any private database producer whose database includes a substantial amount of government-generated data to note that fact with reasonably sufficient details about the government source of the data. By this, we mean, for example, "This database was compiled with substantial amounts of data from the National Weather Service, National Oceanic and Atmospheric Administration, Department of Commerce, Washington, D.C." but not "This database was compiled with information from the Department of Defense." In other words, the disclosure should reasonably direct the user to the government source. Defendants could be given an express defense where the database producer has included substantial amounts of government-generated information and failed to make such a disclosure.

We believe that such a requirement (and defense) would eliminate some *apparent* sole source

situations by pointing the database user to alternative sources for the information. If the worth of the database producer's product was truly in the "value-added," consumers would stay with the private product. Such disclosures might also give government agencies a stronger incentive to maintain the raw data and keep it available to citizens, thus eliminating at least some sole source situations. Generally, we are hopeful that the digital environment and the Internet will, over time, make it possible for government agencies to provide more government-generated information at less cost through public channels.

D. Fourth Principle -- Avoid unintended consequences

Any database misappropriation regime must carefully define and describe the protected interests and prohibited activities, so as to avoid unintended consequences; legislation should not affect established contractual relationships and should apply only prospectively and with reasonable notice

1. Prior contractual relationships

The Administration believes that any database protection law should expressly state that its provisions may not be used to enlarge or limit any rights, obligations, remedies, or practices under agreements entered into prior to the effective date of the law. This is especially important because today, many, if not most, commercially valuable databases are licensed rather than sold. The purpose of such statutory language would be to avoid unbalancing the contractual relationships that have been freely entered into before a database protection bill becomes law. This is a matter of notice and fairness. Providers of databases should not be permitted to assert limitations on use not contemplated at the time of the contract. Similarly, neither database users nor those under contract to produce databases should be able to take unfair advantage of a change in the law to assert rights where existing contracts (including government grants, contracts, or other agreements) may be silent.

2. Prospective Application

We agree wholeheartedly that there should be no liability for conduct prior to the statute's effective date. With respect to situations in which the investment in the database occurred prior to the law's effective date, the situation is more complex. Based on a strict economic analysis, coverage of such databases is not necessary -- the investment occurred without the legal protection. On the other hand, there is some, albeit uncertain, legal protection now. Some incentive still exists deriving from copyright's limited protection, what people still believe to be copyright protection, and by state law. On balance, and especially in the context of a misappropriation approach, we believe that section 4 of H.R. 354 takes an appropriate approach toward this issue.

3. The term of protection

Advocates of database protection have proposed database protection terms of up to 25 years. Alternative views have ranged from criticizing 15 years as too long to the minimalist bill's proposal for more limited rights of unlimited duration. The Administration currently believes that there is no single, optimal term of protection for the wide range of products subject to protection as "databases" or "collections of information." [\(24\)](#)

In the absence of strong indicators of the optimal term for an *ex ante* incentive structure, we believe there are two virtues to the 15-year term of protection. First, it corresponds to the term of protection established in the European Union's Database Directive; this may facilitate emergence of an international standard while allowing us to concentrate on important issues like permitted uses and the flow of government-generated data. Second, we believe that 10-15 years roughly coincides with a substantial number of data producers beginning to maintain their records in digital formats. The presence of such digital archives of raw data is important in helping to avoid as many sole-source situations as possible.

Finally, the Administration would be troubled by any efforts -- present or future -- to establish a term of protection exceeding 15 years. While we recognize that there are and will be some data products which have substantial value after 15 years, the purpose of database protection legislation is to provide an incentive for the creation of new databases; we are doubtful that there are or will be many databases developed with a cost-recovery business plan going beyond 15 years.

4. The "perpetual protection" problem

Some critics of database protection have claimed that while proposals like H.R. 354 call for a fixed term of protection (15 years in this case), they actually raise the specter of "perpetual" protection for non-copyrighted databases. We believe that this is a serious issue that requires careful consideration. The critics' concern about "perpetual protection" has two foundations.

a. "Perpetual protection" from "maintaining": the problem with the "organizing" and "maintaining" criteria

The first source of concern is the word "maintaining" in the basic prohibition. By including "maintaining" as a ground for protection, some database producers may assert that simply maintaining data collected long ago qualifies that data for continuing protection. H.R. 354 seeks to address this problem with the following provision that differs from H.R. 354's predecessor, H.R. 2652, in the bolded text:

"1408(c) Additional Limitation - No 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 15 years after the **portion of the collection that was 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** ~~that is extracted or used~~. **In no case shall any protection under this chapter resulting from a substantial investment of resources in maintaining a pre-existing collection prevent any use or extraction of information from a copy of the pre-existing collection after the 15 years has expired with respect to the portion of that pre-existing collection that is so used or extracted, and no liability under this chapter shall thereafter attach to such acts or use or extraction.**"

The final sentence of section 1408(c) apparently is intended to eliminate the possibility of "maintenance" being used to perpetuate protection for data entries.

The Administration agrees with Chairman Coble that this potential problem must be addressed and appreciates the effort to respond to it. We are concerned, however, that this approach is too complex. We believe that a simpler, more predictable legal schema would be produced by eliminating "maintaining" as a ground for protection in the basic prohibition. In fact, we urge the Subcommittee to consider whether either "maintaining" or "organizing" is needed as an event triggering protection under the statute. We believe that substituting "collecting" for "gathering" and making it the sole basis for protected investment would address this perpetual protection issue and better focus the statute.

The present legislation is motivated by the need to correct the loss of protection for "industrious collection" under the "sweat of the brow" doctrine. Adding protection for "organizing" and "maintaining" would expand the protected investment well beyond what was historically allowed by the courts that embraced that doctrine. The *Warren Publishing* and similar cases involve collecting in the traditional sense, while there is no history or definition for "organizing" or "maintaining." Some aspects of maintaining data such as checking and adding facts are really aspects of "collecting" and should be recognized as such. We also believe that "collecting" data captures much of the value in "organizing" data that can be lost to free-riders. Organizing that involves selection or judgment is protectable under copyright law, even after the decision in *Feist*. Therefore, inclusion of that term here is not necessary to provide an incentive for such activities. On the other hand, merely mounting a database on a server is

part of maintaining it, but mounting data for access does not suffer from the free-riding problem of collecting (i.e., it is an expense that must be borne by the misappropriator as well as the original publisher). For all these reasons, we think it necessary to protect only "collecting."

b. Concern for *de facto* "perpetual protection"

We also believe that there is a potential "perpetual protection" problem that is more complicated. This problem is rooted in the need to provide some type of protection for *revisions* of databases. Legislation that provided protection to new databases but not to revisions of databases, would skew investment. There would be a disincentive to revise proven, useful databases in favor of creating new databases. Reassembling (largely) the same information in a new database would be inefficient not only for data gatherers, but for data users who -- in order to use the most current data -- would have to accustom themselves to the format of the new database. Therefore, any database protection legislation should offer protection for revisions of existing databases, so that new iterations of a protected database are themselves protected. But this means that eventually there may be unprotected data entries (from iterations of the database older than 15 years) intermingled with protected data entries (from more recent iterations).⁽²⁵⁾

This gives rise to a potential problem. In the classic case of a copyrighted book, the text loses protection at the end of its term, although new, revised versions of the text may enjoy fresh periods of protection. This means that one can find unprotected texts of *The Raven* or *Leaves of Grass* in libraries all over the country. At the same time, new versions of these books can be under some copyright protection (including new introductions, abridgements, "notes," artwork, etc.). It is possible to compare the two versions -- old, unprotected and new, protected -- side-by-side.

In the digital, on-line environment, content producers may choose not to sell copies of their works; access to a database may instead be licensed to users. The advantage is that the database user can receive the most current version of the compilation. The disadvantage is that the user may lack access to an old version of the database to compare old and new entries. The question is, how can a user, accessing only the newest version of a database that has gone through many iterations, distinguish unprotected data entries from protected data entries? In Appendix A we give a simple example of how this problem would arise.

While the Administration believes that the new language of section 1408(c) helps ensure that the bill provides no *de jure* perpetual protection, we remain concerned that the digital environment could produce *de facto* perpetual protection because users would be unable to distinguish protected and unprotected data and, therefore, would be chilled in their use of unprotected data.⁽²⁶⁾ Such inadvertent extension of the protection afforded by H.R. 354 could exacerbate other concerns, including the "sole source" issue and the constitutionality of the law.

There have been varied proposals to address this problem. One proposal has been to "tag" data entries so that older, unprotected data can be distinguished from protected data. We are not in a position to comment on the feasibility, whether technological or economical, of this suggestion. Another proposal - which is set out in the Senate discussion draft - would be the establishment of a deposit system to ensure that older, unprotected versions of the databases would be available to the public. We believe that the storage demands of such a deposit system would exceed anything the Copyright Office or the Patent and Trademark Office now handles. It is also not clear how the costs of such a deposit system should be apportioned.

At different junctures in this statement, we have recommended establishing express statutory defenses to remedy possible problems in a database protection; we make the same type of suggestion here. Where the database that is the subject of a litigation is the descendant of a now unprotected database and has substantial elements in common with that unprotected database, the defendant should be able to raise, as a defense, that the most recent unprotected iteration of the database is not reasonably

publicly available.

In other words, if Smith Industries has been issuing the "Smith Industrial Database" annually since 1980, and then in 1999 if Smith Industries sues someone for unauthorized distribution of the "1999 Smith Industrial Database" the defendant can raise as a defense that the 1983 Smith Industrial Database is no longer reasonably publicly available. If the 1983 database is reasonably publicly available, there is no such defense.

The virtue in this approach in comparison to mandatory "tagging" or deposit systems is that it allows each private enterprise to determine *how* to make its now unprotected database available to the public. Moreover, the database producer does not have to make this final decision until the term of protection is over. Some concern has been expressed about this proposal by database producers who produce continuously updated databases; their situation in relation to this proposed defense merits examination. But, as we said above, we propose the defense when the protected database "is the descendant of a now unprotected database and *has substantial elements in common with that unprotected database.*" We believe that for many continuously updated databases, the most recent database would have almost no elements in common with their 15-year ancestor.[\(27\)](#)

5. The "sole source" problem

There has been much discussion of what is called the sole-source problem: that many markets for data will be supplied by only one database provider. The sole-source problem arises most acutely when one entity controls access to a unique, unreplicable collection of information, such as weather data that occurs once and cannot be replicated. This control may arise either purposefully, as with an exclusive contract with the data's original generator, or incidentally, when the data's original generator ceases to maintain it. Other practical sole-source situations can arise when an existing database operates as a natural monopoly; that is, it is possible, but not economically efficient, for someone else to build the dataset independently.

Even now, a sole-source may use contracts to preserve its market position against free riding by would-be competitors. Any form of database protection carries with it the possibility that it could further insulate a sole-source database provider against potential competition. Consequently, it will be important that any database protection legislation incorporate provisions that guard against the possibility that sole-source database providers will employ their new rights to the detriment of competition in related markets.

A partial answer to the sole-source problem is a savings clause such as the one in H.R. 354, providing that nothing in the bill operates to the detriment of federal antitrust law. Thus, for example, database owners would be as subject as any other economic actors to the application of the essential facilities doctrine, which prohibits owners of assets that are essential to the ability to compete in a market, and are not feasible to replicate, from refusing to deal with firms that need that access. On the other hand, this doctrine has been invoked relatively rarely, and understandably so: part of the incentive for the development of any valuable product or service is the hope that the product or service will be so attractive to consumers that it will become dominant. Regularly compelling access to valuable products and services could diminish their developers' incentives to invest in them in the first place.

At the same time, in markets such as data collection and dissemination, where natural-monopoly characteristics suggest that consumer choice among competing database products and services will not be common, some safety valve over and above the rarely used essential-facility doctrine may be necessary to ensure that database providers are not able to deny access to firms that require it in order to compete in downstream markets. Additional possibilities include the development of doctrines comparable to the misuse doctrine in patent and copyright law or, in extreme cases, the idea/expression merger doctrine in copyright law.[\(28\)](#)

As with some other problems we have identified above, however, much of the concern arising from the sole-source problem can be eased by defining both the protected activity and the prohibited conduct narrowly. If the bill protects only data collection and generation, it will be covering value-adding conduct that enhances welfare, even though only one firm may find it worthwhile to engage in collecting and disseminating a particular type of database. Similarly, to the extent that the bill prohibits only distribution and extraction for the purpose of distribution, while conversely permitting transformative uses of data, it would leave data providers free to add value and enter markets that the original data collector's work alone was incapable of serving.

E. Fifth Principle -- Balance protection with permitted uses

Any database misappropriation regime should provide exceptions analogous to fair use principles of copyright law; in particular, any effects on non-commercial research should be *de minimis*.

Given the difficulty of foreseeing how "substantiality," "extraction" and other legislative terms will play out in a complex and rapidly changing environment, we expressed concern last summer that H.R. 2652 lacked a balancing mechanism analogous to the fair use doctrine in copyright sufficient to address the wide range of circumstances in which information is aggregated, used, and reused. We were especially concerned that the section 1203(d) exception for non-commercial research and educational uses did not ensure against disruption of legitimate non-commercial research, and that educational activities were not disrupted by the prohibition against commercial misappropriation. Last year, we also were concerned with equitable issues of access and use that may be especially important in markets exclusively served by a single data producer.

In reviewing the permitted acts provisions of H.R. 354 (section 1403), we would like to suggest, as an initial matter, that the Subcommittee rearrange the various "permitted acts" to move more clearly from *absolutely* shielded activities for *all persons* (such as use of insubstantial parts (1403(b)) to the more limited shields on activities set out in 1403(a)(2). We propose that the Subcommittee reorder section 1403 as shown in Appendix B. We believe that this reordering would provide legislation that is easier to understand and a clearer platform for full discussions on whether the permitted activities adequately address policy and constitutional concerns. This proposed reordering is separate from any substantive recommendations.

As to the substantive elements of the permitted acts section, the Administration is pleased that H.R. 354 limits the liability for nonprofit educational, scientific, and research purposes to uses that harm directly the actual market and that the legislation now includes as section 1403(a)(2)(A) provisions for "additional reasonable uses" similar to the fair use provisions of section 107 of the Copyright Act. However, we are concerned that the last sentence in section 1403(a)(2)(A) could be interpreted as overriding the criteria in section 1403(a)(2)(A) with a standard that differs in form but not in substance from the basic operating provisions of section 1402. The Administration would not agree with any intent to override a "fair use"-like balancing test; on the other hand, if the last sentence of 1403(a)(2)(A) is intended only to restate the basic prohibition without disturbing the balancing test, it is extraneous language. We therefore recommend its deletion.

We recognize the desire to avoid the precise fair use terminology of the Copyright Act in order to make clear that the legislation is grounded in misappropriation rather than intellectual property. However, in the interests of transparency and predictability, we believe that the fair use principles of copyright are a sound platform on which to build. Providing the safeguard of familiar fair use criteria can help minimize any unintended consequences of the untested basic operating provisions of section 1402. We believe that this would give courts the tools they need to do justice in particular situations.

The fair use factors may need to be framed or supplemented to allow courts to take into account that the subject matter is industrious collection rather than original expression, that the protected interest is purely economic, and that the proscribed behavior is a form of unfair competition. The provision

would also have to be recrafted to focus on distribution, rather than use, if the basic prohibition were amended as we have suggested above. Courts might also be called on to recognize the unique conditions of some database markets. But we believe that the vast experience of courts in using the judicially-crafted principles of fair use should be built into database protection legislation. It is worth noting that in the 23 years since Congress codified the fair use factors, it has neither narrowed nor expanded these factors. While it may be appropriate to diverge from copyright fair use in creating the permitted uses regime for database legislation, the differences between the two should be clearly understood and recognized by concerned parties.

Finally, we would reiterate a point made earlier: the scope of the basic prohibition will determine the weight that the permitted uses section must bear in judging both the policy and constitutionality of any database protection legislation.

F. Sixth Principle -- Ensure protection for U.S. companies abroad and promote harmonization

Consistent with the goals of the World Trade Organization (WTO) and U.S. trade policy, legislation should aim to ensure that U.S. companies enjoy available protection for their database products in other countries on the same terms as enjoyed by nationals of those countries.

There has been some discussion in the United States about the effects of the European Union's 1996 Database Directive (EU Directive) on American database producers. The EU Directive requires European Union Member States to provide *sui generis* protection for databases, but denies this protection to nationals of any foreign country unless that country offers "comparable protection to databases produced" by EU nationals.⁽²⁹⁾

The Administration opposes such "reciprocity" requirements, both domestically and internationally. We believe that commercial laws (including intellectual property and unfair business practices laws) should be administered on national treatment terms, that is, a country's domestic laws should treat a foreign national like one of the country's citizens. This principle is embodied in Article 3 of the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS Agreement) as well as more generally in the Paris Convention for the Protection of Industrial Property and the Berne Convention for the Protection of Literary and Artistic Works.

The Administration believes that Congress should craft U.S. database protection legislation to meet the needs of the American economy. A database protection law properly balanced for the robust digital economy of the United States will serve as a model for other countries that hope to build businesses, employment, and economic activity in the new millennium.

At the same time, we believe that a law along the lines of H.R. 354 (with proper attention to the concerns we have identified) will amply provide protection "comparable" to that provided by national laws implementing the EU Directive. From the perspective of a private database producer, a misappropriation law as discussed in both the last and current Congress would, we believe, provide a cause of action and meaningful remedies in the same range of situations in which the laws implementing the EU Directive provide a cause of action and meaningful remedies.⁽³⁰⁾

Although we believe that a law along the lines of H.R. 354 would provide American database makers with protection under the EU Directive's reciprocity provision, the Administration would, for the reasons stated above, oppose any effort to put automatic reciprocity provisions into American law in this area. United States Trade Representative Charlene Barshefsky cited the reciprocity provision of the EU Directive as a subject of concern in announcing the Administration's 1998 Special 301 Review.

While we believe that a United States database protection law should adhere to a national

treatment model, the Administration would support an appropriately crafted provision that would allow the President to affirmatively deny database protection to foreign nationals on the appropriate finding by Executive Branch agencies such as the USTR and/or the Department of Commerce. This could, for example, be achieved by statutory language or legislative history making database protection for foreign nationals subject to USTR's Special 301 process.

G. Additional Issues

1. Gradations of Criminal Liability

While we agree with Chairman Coble's decision to shield non-profit researchers and educators from any criminal liability under section 1407, we believe that the existing criminal provisions should be further refined, particularly by drawing a distinction between misdemeanor and felony conduct and requiring minimum amounts of damage under each. This will expand the range of charging options available to prosecutors. We have attached our recommendation for statutory language as Appendix C.

2. Data-Gathering Activities of Law Enforcement Agencies

We believe it is important to make clear that the legitimate data-gathering activities of law enforcement and intelligence agencies will not be affected by the bill. While we believe that intelligence gathering and national security activities are already shielded from liability by section 1402 in that these activities will not cause "harm to the actual or potential" market of the product, we propose an additional statutory provision and legislative history as shown in Appendix D to confirm that these activities fall outside the bill's reach.

3. Administration Study

The Administration believes that, given our limited understanding of the future digital environment and the evolving markets for information, it would be desirable to conduct an interagency review of the law's impact at periodic intervals following implementation of the law. Such a government study might be conducted jointly by the Department of Commerce, the Office of Science and Technology Policy, and the Department of Justice in consultation with the Register of Copyrights and other parties. We believe that such a study should not be limited to any one set of issues or concepts; rather, it should explore issues including: database pricing before and after enactment of the law; database development before and after enactment of the law; international protection for American database producers; the impact of the law on scientific research and education; access issues; and "sole source" databases.

I thank the Subcommittee for the opportunity to appear before you today and look forward to working with you during the legislative process. I would be pleased to answer any questions that you may have at this time.

APPENDIX A

Imagine that in 2000, a database producer makes a database; we will designate the first twelve entries alphabetically:

A

B
C
D
E
F
G
H
I
J
K
L

In 2003, it "expands and refreshes" the database, so that the first fifteen entries are as follows:

A
B
BB
C
D
E
F
FF
G
H
I
J
K

KK

L

In theory, under H.R. 354 in the year 2016, all of the entries except BB, FF, and KK lose protection -- and can be copied in their entirety. The problem is that if the database is provided via on-line services, there may be no means for the user to know which entries are unprotected because they were original entries and which entries are protected because they are the result of maintenance investment within the past 15 years. One commentator has suggested that new entries be electronically "tagged," so that a user can readily determine what is protected and what is not, i.e.

A

B

BB

C

D

E

F

FF

G

H

I

J

K

KK

L

Another possible solution would be to require any database producer that wanted to enjoy protection for a revision of their database after the fifteen year period to make (or have made) the original, no longer-protected database available in a reasonable format. This would be the electronic equivalent of the

old copy of *Wuthering Heights* in the public library. The original database need not be *as* available as the new version -- just as old library books usually are not as available as books at retail stores, but it should reach some standard of public access.

Appendix B

H.R. 354

PROPOSED RE-WORDING OF THE PROVISIONS ANALOGOUS TO FAIR USE IN COPYRIGHT LAW

(with minimal edits)

Sec. 1403. Permitted Acts **and Uses**

(a) GATHERING OR USE OF INFORMATION OBTAINED THROUGH OTHER MEANS-

Nothing 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.

(b) INDIVIDUAL ITEMS OF INFORMATION AND OTHER INSUBSTANTIAL PARTS-

Nothing in this chapter shall prevent the extraction or use of an individual item of information, or other insubstantial part of a collection of information, in itself. An individual item of information, including a work of authorship, shall not itself be considered a substantial part of a collection of information under section 1402. Nothing in this subsection shall permit the repeated or systematic extraction or use of individual items or insubstantial parts of a collection of information so as to circumvent the prohibition contained in section 1402.

(c) USE OF INFORMATION FOR VERIFICATION-

Nothing in this chapter shall restrict any person from extracting or using a collection of information within any entity or organization, for the sole purpose of verifying the accuracy of information independently

gathered, organized, or maintained by that person. Under no circumstances shall the information so used be extracted from the original collection and made available to others in a manner that harms the actual or potential market for the collection of information from which it is extracted or used.

(d) **NEWS REPORTING-**

Nothing in this chapter shall restrict any person from extracting or using information for the sole purpose of news reporting, including news gathering, dissemination, and comment, unless the information so extracted or used is time sensitive and has been gathered by a news reporting entity, and the extraction or use is part of a consistent pattern engaged in for the purpose of direct competition.

(e) **TRANSFER OF COPY-**

Nothing in this chapter shall restrict the owner of a particular lawfully made copy of all or part of a collection of information from selling or otherwise disposing of the possession of that copy.

Sec. 1404. Additional Reasonable Uses

(a) **CERTAIN NONPROFIT EDUCATIONAL, SCIENTIFIC, OR RESEARCH USES-**

Notwithstanding section 1402, no person shall be restricted from extracting or using information for nonprofit educational, scientific, or research purposes in a manner that does not harm directly the actual market for the product or service referred to in section 1402.

(b) **GENERAL REASONABLE USES-**

Notwithstanding section 1402, 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 reasonable under the circumstances. In determining whether such a reasonable under the circumstances, the following factors shall be considered:

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

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.

~~(B) DEFINITION- 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.~~

Renumber sections 1404 and subsequent

Add to:

Sec. 1401. Definitions

(5) INDIVIDUAL ACT -- 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.

Appendix C

§ 1407. Criminal offenses and penalties

(a) Violation.--

(1) In General.- Any person who violates section 1202 willfully either

(A) for purposes of direct or indirect commercial advantage or financial gain, or

(B) causes loss or damage aggregating \$100,000 or more during any 1-year period to the person who gathered, organized, or maintained the information concerned, or

(C) causes loss or damage aggregating \$50,000 or more in any 1-year period to the person who gathered, organized, or maintained the information concerned,

shall be punished as provided in subsection (b).

(2) Inapplicability. --This section shall not apply to any employee or agent of a nonprofit educational, scientific, or research institution, library, archives, or law enforcement agency acting within the scope of his or her employment.

(b) Penalties.--

(1) Any person who commits an offense under subsection (a)(1)(A) shall be fined not more than \$250,000 or imprisoned for not more than 5 years, or both;

(2) Any person who commits a second or subsequent offense under subsection (a)(1)(A) shall be fined not more than \$500,000 or imprisoned for not more than 10 years, or both;

(3) Any person who commits an offense under subsection (a)(1)(B) shall be fined not more than \$250,000 or imprisoned for not more than 3 years, or both;

(4) Any person who commits a second or subsequent offense under subsection (a)(1)(B) shall be fined not more than \$500,000 or imprisoned not more than 6 years, or both;

(5) Any person who commits an offense under subsection (a)(1)(C) shall be fined not more than \$100,000 or imprisoned not more than 1 year, or both.

(c) Victim Impact Statement.--

(1) During preparation of the presentence report pursuant to Rule 32(c) of the Federal Rules of Criminal Procedure, victims of the offense shall be permitted to submit, and the probation officer shall receive, a victim impact statement that identifies the victim of the offense and the extent and scope of the injury and loss suffered by the victim, including the estimated economic impact of the offense on that victim.

(2) Persons permitted to submit victim impact statements shall include--

(A) persons who gathered, organized, or maintained the information affected by conduct involved in the offense; and

(B) the legal representatives of such persons.

Appendix D

Possible Addition to §1403 to Address

National Security/Intelligence Concerns

Addition of new subsection (g) to §1403 Permitted Acts:

"(g) Nothing in this chapter shall prohibit an officer, agent, or employee of the United States, a State, or a political subdivision of a State or a person acting under contract of one of the enumerated officers, agents, or employees from extracting and using information as part of lawfully authorized investigative, protective, or intelligence activities."

Proposed Legislative History to Accompany §1403(g):

Intelligence gathering and national security activities are already shielded from liability by section 1402 in

that these activities will not cause "harm to the actual or potential" market of the product. Section 1403 (g) is offered to further clarify and confirm that these activities and law enforcement activities fall outside the bill's reach. Subsection 1403(g) is not intended to permit law enforcement or intelligence agencies to use commercially available databases without liability where the use occurs in normal ministerial functions or publicly-known activities of the agency, if such use would cause harm to the market as detailed in section 1402. For example, section 1403 (g) would apply to covert or undercover investigative or intelligence activities where the officer, agent, or employee may be called upon to access databases - - physically or through computer networks - - without the knowledge of the database producer or the owner (or license holder) of that copy of the database. Section 1403 (g) helps make it clear beyond any doubt that law enforcement and intelligence agencies can continue to conduct any lawfully authorized activities without becoming liable under this Act.

NOTES

1. *Feist Publications v. Rural Telephone Service Corp.*, 499 U.S. 340 (1991).
2. Including the National Oceanic and Atmospheric Administration (NOAA), the National Institute of Standards and Technology (NIST), the U.S. Geological Survey, the Department of Energy, and the PTO.
3. See, e.g., National Research Council, *Bits of Power* (1997) at 135; U.S. Patent and Trademark Office, *Report on and Recommendations from April 1998 Conference on Database Protection* (1998) at 4-7; Letter from Federal Trade Commission Chairman Robert Pitofsky to Congressman Tom Bliley, September 28, 1998 at 6-7. See also Institute of Intellectual Property, Tokyo, Japan, *Database Protection on the Borderline of Copyright Law and Industrial Property Law* 5 (1998); Wendy Gordon, *Asymmetrical Market Failure and Prisoner's Dilemma in Intellectual Property*, 17 U. Dayton L. Rev. 853, 863-865 (1992) (describing conditions when additional protection is needed); Dan L. Burke, *The Market for Digital Piracy*, in Brian Kahin and Charles Nesson, eds., *Borders in Cyberspace* (1997), 205 (describing databases on the Internet as classic "public good" problem that may require special law); J.H. Reichman and Pamela Samuelson, *Intellectual Property Rights in Data?*, 50 Vanderbilt L. Rev. 51, 55 (1997) (critical of EU Database Directive and H.R. 3531, but recognizing that risks of market failure may keep data production at "suboptimal levels"); M. Powell, *The European Database Directive: An International Antidote to the Side Effects of Feist?* 20 Fordham International L. J. 1215, 1250 (1997).
4. There has been much discussion among commentators about the differences between a *sui generis* form of protection as was proposed in H.R. 3531 in the 104th Congress and the "misappropriation" approach proposed in the present H.R. 354. The Administration believes that the misappropriation theory provides an appropriate model for database protection in American law. The United States has substantial case law on the misappropriation of information as a form of unfair competition which should help courts interpret any database protection law built on a misappropriation model. Placing database protection in the framework of unfair competition will also allow courts and commentators to draw appropriately from the rich body of cases in trademark law and unfair business practices.

The Administration believes that any treaty on database protection that emerges from

on-going discussions at the World Intellectual Property Organization should permit each treaty signatory to provide any mandated database property protection through the legal mechanism most appropriate to its domestic law, whether through misappropriation, *sui generis* protection, or a simple extension of their domestic copyright and neighboring rights laws. The critical issue is not the legal framework used, but whether the law provides private citizens with comparable rights to protect their investments in different jurisdictions.

5. In contrast, the basic prohibition in what some have called the "minimalist" proposal put forward by some database users seems too narrow as a policy matter. See Section 1401 of "Proposed Bill to Amend Title 17, United States Code, To Promote Research and Fair Competition in the Databases Industry," Statement by Senator Orrin Hatch, *Congressional Record*, January 19, 1999, at S320. This minimalist basic prohibition appears to bar only misappropriation of an entire database, but to permit appropriation of a large percentage of the same database, even for a commercial purpose in competition with the database creator. There are also constitutional concerns with the minimalist approach, albeit not as serious as with H.R. 354.

6. *Warren Publishing v. Microdos Data Inc.*, 115 F.3d 1509 (11th Cir. 1997) (en banc) cert. denied 118 S.Ct. 197 (1997).

7. 18 U.S.C. § 1030 would appear to create *some* criminal liability for database misappropriation by individuals in the on-line environment. Subsection 1030(a) (2) (C) creates criminal liability when a person "intentionally accesses a computer . . . and thereby obtains . . . information from an protected computer if the conduct involved an interstate or foreign communication," while 1030(a)(4) creates criminal liability when a person "knowingly and with intent to defraud, accesses a protected computer without authorization . . . and by means of such conduct . . . obtains anything of value" in excess of \$5,000. We assume that the server holding a commercial database would fall within the definition of a "protected computer" because it would be "a computer . . . which is used in interstate or foreign commerce or communication [1030(e) (2)(B)]. Subsection 1030(g) also creates civil liability where there has been a "violation" of the section.

8. Substantial harm is a familiar standard applied by courts in a variety of circumstances. See, e.g., *Gulf & Western Industries, Inc. v. United States*, 615 F.2d 527 (D.C. Cir 1979) (enunciating standard for when disclosure of commercial information in government's possession would cause substantial harm to competitive position of private firm); *Miami Herald v. SBA*, 670 F.2d 610 (5th Cir. 1982) (same standard); *Simmons v. Diamond Shamrock Corp.*, 844 F.2d 517 (8th Cir. 1988) (determining whether failure to comply with ERISA reporting and disclosure requirements caused substantial harm); *Warner Bros. v. U.S. ITC*, 787 F.2d 562 (Fed. Cir. 1986) (ITC temporary exclusion order depends on showing of immediate and substantial harm in the absence of such relief); *Olson v. Stotts*, 9 F.3d 1475 (10th Cir. 1993) (substantial harm standard used for liability in delay in medical care) . . . not to mention the use of "substantial harm" as a standard in preliminary injunction cases. See, e.g., *N.A.A.C.P. v. City of Mansfield*, 866 F.2d 162, 166 (6th Cir. 1988).

9. 105 F.3d 841, 852 (2d Cir. 1997)

10. Evidence that the defendant had not diminished the plaintiff's incentive to produce the database could, however, be the same type of evidence that shows how "transformative" the defendant's product is and how far its sales are from the original product's market. In this way, the same evidence might enter a litigation under an appropriately broad "permitted uses" section.

11. A Framework for Global Electronic Commerce is available at:
<http://www.ecommerce.gov/framework.htm>.

12. Conceptually, a distinction can be drawn between data "gathered" and data "generated." The decennial Census *gathers* data about Americans; the Hubble telescope gathers astrophysical data by capturing images of events that have already occurred in distant parts of the Universe. In contrast, when the U.S. Government established the "zip code" system, it generated data that did not exist before. The government *generates* data in the form of new legal opinions, new tax tables, new databases of each day's recipients of Medicare or Medicaid payments. We refer to the results of all these activities, including research funded by the government through grants or contracts, as "government-generated data" or "government-funded data."

13. Office of Management and Budget Circular A-130 Revised [Section 7.b, "Basic Considerations and Assumptions"], available at: <http://www.whitehouse.gov/omb/circulars/a130/a130.html>, *hereinafter* "Circular A-130".

14. The U.S. Government's position on the importance of the free exchange of such data has been stated often, including in the "Bromley Statement" on climate change information. *See* Data Management Global Change Research Policy Statement, Office of Science and Technology Policy, The White House, July 2, 1991.

15. Circular A-130, Appendix IV "Analysis of Key Sections," section 3 "Analysis."

16. "Agencies must justify the creation or collection of information based on their statutory functions. Policy statement 8a(2) uses the justification standard -- 'necessary for the proper performance of the function of the agency' -- established by the [Paperwork Reduction Act] (44 U.S.C. § 3508)." Circular A-130, Appendix IV "Analysis of Key Sections," section 3 "Analysis."

17. This distinction would apply to more than universities. Many government agencies offer their unique capabilities to the private sector on a reimbursable basis. At the Department of Energy, for example, these transactions can be Cooperative Research And Development Agreements (CRADAs) which are "100% funds-in" agreements or "Work for Others" agreements or User Faculty agreement: that is, the private entity provides 100% of the operating funds for the research which is conducted at a government laboratory. We believe that these privately funded research projects could reasonably give rise to collections of information protectable under a database protection law *because* in judging the equities of the relative contributions to the final database product, there is little or no government investment. Failure to provide protection in such cases would discourage businesses from entering into these agreements. This would sharply curtail the ability of the government to enhance the competitiveness of the private sector.

18. Section 1301(6)(B), *Congressional Record*, January 19, 1999, at S322.

19. For example, the Agriculture Department works under a directive to "diffuse among people of the United States, useful information on subjects connected with agriculture . . ." 7 U.S.C. § 2201, . . . while NASA has a mandate to "provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof," 42 U.S.C. section 2473(a)(3) and 42 U.S.C. section 2051 requires the Department of Energy to insure the continued conduct of research and prohibits "any provisions or conditions [on research] which prevent the dissemination of scientific or technical information . . ." 42 U.S.C. § 2051 (d). Statutes such as the Freedom of Information Act and the Government in the Sunshine Act "establish a broad and general obligation on the part of Federal agencies to make government information available to the public and to avoid erecting barriers that impede public access." Circular A-130 , Appendix IV "Analysis of Key Sections," section 3 "Analysis." Other departments and programs are under express regulatory mandates to make compilations of information available to the public. For example, in some of their mapping and surveying programs, the Departments of the Interior and Commerce are under a mandate to provide data products "in a format that can be

shared with other Federal agencies and non-Federal users." Office of Management and Budget, Circular A-16 Revised [*Coordination of Surveying, Mapping, and Related Spatial Data Activities*], section 2.

20. This same balance was expressed by Weiss and Backlund as follows: "On the one hand, this means that the Government should not try to duplicate value-added information products produced by the private sector. On the other hand, it means that the government should actively disseminate its information - particularly the raw content from which value-added products are created - at cost and not attempt to exert copyright-like controls or restrictions." Peter N. Weiss and Peter Backlund, *International Information Policy in Conflict: Open and Unrestricted Access versus Government Commercialization*, in Brian Kahin and Charles Nesson, eds., *Borders in Cyberspace* (1997), 300, 303.

21. A disclaimer capturing the spirit of this requirement is that found in the U.S. Industry and Trade Outlook (1998) published by McGraw-Hill in cooperation with the Department of Commerce. The disclaimer states: "Portions of this publication contain work prepared by officers and the employees of the United States Government as part of such person's official duties. No copyright is claimed as to any chapter or section whose designated author is an employee of the United States Government, except that copyright is claimed as to tables, graphs, maps or charts in any chapters or sections of this publication if the sole designated source is other than the United States Government."

22. The question of databases of court opinions is complicated by the fact that there are arguably two sets of data intertwined in a commercial volume of court opinions. First, there is the publicly-generated opinions. Second, there is the privately-generated elements, including the pagination of the volume. In *Matthew Bender v. West Publishing*, 158 F.3d 674, 1998 U.S. App. LEXIS 30790, 48 U.S.P.Q.2d (BNA) 1560, (November 3, 1998), the Second Circuit recently concluded that the pagination in privately published court volumes is non-copyrightable. It appears that H.R. 354 would allow second publishers to note where text starts and stops on different pages as independently observable facts under section 1403(c) of the bill.

23. Statutory requirements of mandatory disclosure of government funded research or government collected information may impinge upon the government's legal and moral obligations to shield some forms of data from disclosure, e.g., private personal data collected in medical research or proprietary business data shared with the government on the condition of non-disclosure.

24. This is similar to economists' efforts to establish the optimal term of protection for copyrighted works where, for example, copyrighted software has a much shorter product cycle than copyrighted books and films which retain significant commercial value for decades.

25. The legislative history for H.R. 2652 in the last Congress also bore on this issue, stating:

"[N]o action can be maintained more than fifteen years after the investment of resources that qualified that portion of the collection of information that is extracted or used. This language means that new investments in an existing collection, if they are substantial enough to be worthy of protection, will themselves be able to be protected, ensuring that producers have the incentive to make such investment in expanding and refreshing their collections. At the same time, however, protection cannot be perpetual; the substantial investment that is protected under the Act cannot be protected for more than fifteen years. By focusing on that investment that made the particular portion of the collection that has been extracted or eligible for protection, the provision avoids providing on-going protection to the entire collection every time there is an additional substantial investment in its scope or maintenance." (Legislative Report)

26. At the same time, we believe that this potential problem arises with particular kinds of databases. Some databases are revised extensively and constantly; for these databases, the value of the database is much

shorter than 10 or 15 years. Stock exchange price listings are the most extreme example, but other lists -- realtors' sale listings and used car valuations also fall in this category. Other databases will be revised rarely once a definitive version is completed, *i.e.* a database of Union warships in the Civil War or the passengers on the *Mayflower*. The databases for which the "perpetual protection" problem arises are between these extremes: they are databases that have value over many years and require substantial, but not total, revision. Examples might include a historical database of the batting statistics of all baseball players in the major leagues or pharmaceutical or toxicological databases used in the medical professions.

27. We recognize that this might still leave the problem of an old, but still protected iteration of the frequently refreshed database having a defense raised against it because the most recent *unprotected* database is not reasonably publicly available. At the same time, we are not convinced that producers of frequently refreshed databases cannot find means to ensure that at least intermittent, historic versions of their databases are reasonably publicly available.

28. Further, a requirement that database providers notify users of their intent to assert rights against misappropriation can mitigate against the possibility of some sole-source situations ever developing; if users are on notice that they may be liable for their conduct involving data from a particular database, they may have reason to seek out alternative sources of the data, so that they will not be locked into a single, dominant source down the road.

29. This is established in Recital 56 of the EU Directive. Recital 56 also provides that a foreign national will enjoy database protection when those "persons have their habitual residence in the territory of the Community." This may provide protection to American database producers who have substantial business operations in EU Member States. Pursuant to Article 11/3 of the EU Directive, a determination whether a foreign state offers "comparable" protection must be made by the European Council based on recommendations from the European Commission.

30. The EU Directive is not a national law. It "directs" the Member States of the EU to implement a legal framework. H.R. 354 would have to be compared, for example, to German, Dutch, and/or Italian law to make the proper comparison of national law to national law. Such a comparison is well beyond the scope of this statement.

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

E-News	Recent	Copyright	Govt. Info	Telecom.	Other Issues	GIS	Letters
------------------------	------------------------	---------------------------	----------------------------	--------------------------	------------------------------	---------------------	-------------------------



© Association of Research Libraries, Washington, DC

Web Design by [Michael Miller](#)

Maintained by [ARL Web Administrator](#)

Last Modified: June 24, 2001