Hearing on the Collections of Information Antipiracy Act

http://www.house.gov/judiciary/106-lede.htm

Statement of

Joshua Lederberg
President-emeritus
Rockefeller University

on behalf of the
National Academy of Sciences
National Academy of Engineering
Institute of Medicine
and
American Association for the Advancement of Science

before the

Committee on the Judiciary
U.S. House of Representatives

18 March 1999

Testimony

Joshua Lederberg
President-emeritus
Rockefeller University

18 March 1999

Introduction

My name is Joshua Lederberg. I have been asked to testify on behalf of the U.S. National Academy of Sciences, National Academy of Engineering, Institute of Medicine (the "Academies"), and the American Association for the Advancement of Science (AAAS). As you know, the three Academies were chartered by Congress to provide advice to the federal government and to the nation on scientific, technical, and medical issues. The AAAS is the umbrella organization for over 250 professional scientific and engineering societies in the United States, with more than 140,000 individual members. I was elected as member of the National Academy of Sciences in 1957 and am a charter member of the Institute of Medicine. I also am Fellow of the AAAS. My biographical summary is attached at the end of this statement.

I am grateful to have the opportunity to testify to you today about H.R. 354, the "Collections of Information Antipiracy Act." This proposed legislation concerns a topic about which the Academies, the AAAS, and indeed the entire research and education community, have an abiding interest and continuing concerns. It also is one that I have had the opportunity to consider from several pertinent perspectives: as a professor and former president at Rockefeller University; as a Nobel prize winning researcher in molecular biology, genetics, and bioinformatics; as an adviser to government science agencies and to scientific publishers; and not least as a creator and user of both commercial and nonprofit scientific databases. I remain an adviser to such enterprises, but I am here formally representing only the Academies and the AAAS. Nevertheless, it is this integration of perspectives and interests from the private sector, government, and academia that I believe is so important to balancing the interests in the pending legislation of both original database creators and providers on the one hand, and of all downstream users on the other.

I would like to make several points in this testimony:

- Access to factual data is essential to furthering our understanding of nature, to medical and technical progress, and to the validation of scientific claims. The essence of the scientific method is relentless critical discourse; without it, the authenticity of knowledge claims rapidly deteriorates. Indeed, a thriving public domain for data, fostered by government policies that guarantee full and open access to data, benefits all downstream users, including the commercial database industry.

- One of the major drivers for new database legislation in the United States is the reciprocity clause of the European Union's Directive on Databases. As discussed below, the Directive imposes strong economic and legal restrictions on the conditions of availability and use of factual data in databases in Europe. The Directive could have adverse consequences not only in Europe, but for cooperative U.S.-E.U. academic endeavors. Before adopting equivalent strong and unprecedented database protection in our country to satisfy a European reciprocity provision of questionable validity, it is essential to consider carefully the underlying rationale and potential impacts to our research and education base and to our economy.

- Significant legal, technical, and self-help protection measures to counter database piracy are already available. There is no evidence of a crisis in the development of new commercial databases. While we support the adoption of new measures that are designed to address specific problems, such as wholesale database piracy, we are opposed to the creation of unprecedented and unjustified new rights in factual information that do not balance the legitimate competing interests.

- Although we appreciate the two changes that were made to H.R. 2652 in H.R. 354, we find the bill as currently proposed to be unacceptable. We would like to draw the CommitteeÌs attention to the

legislative proposal, "The Database Fair Competition and Research Promotion Act of 1999," which we support, as well as to the progress made on achieving a compromise on H.R. 2652 in negotiations sponsored by the Senate Committee on the Judiciary last summer.

- The Academies and the AAAS remain committed to working with Congress on crafting a well reasoned and balanced database protection bill that serves the interests of our nation, and not just one segment of the publishing community.

I also would like to note at the outset that we are introducing into the official record several attachments to this testimony. The first, labelled Appendix A1, A2, and A3, presents an abridged selection of the AcademiesÌ analytical summaries and alternative proposals to Title V of H.R. 2281, the successor bill to H.R. 2652, which were submitted during the Senate Committee on the Judiciary negotiations last summer. The second is a September 4, 1999 letter from Professor Harvey Perlman of the University of Nebraska Law School to Senator Orrin Hatch about unfair competition law.

The Need to Maintain Our Traditional Public Domain for Factual Data

Scientific and engineering research drives our nationÌs progress. Society uses the fruits of such research to expand the worldÌs base of knowledge and applies that knowledge in myriad downstream applications to create new wealth and to enhance the public welfare. Indeed, the policy of the United States has been to support a vibrant research enterprise and to assure that its productivity is exploited for national gain. Thus, freedom of inquiry, the open availability of scientific data, and the open publication of results are cornerstones of our research system that U.S. law and tradition have long upheld.

The consequences of these wise policies has been spectacular. For many decades, the United States has been the leader in the collection and dissemination of scientific and technical data and in the discovery and creation of new knowledge. Our nation has used that knowledge more effectively than any other nation to support new industries and applications, such as the biotechnology industry and the discovery of new diagnostics and cures for hereditary and other diseases.

A necessary component of these past and continuing achievements has been the wide availability of scientific and technical data and information, ranging from raw or minimally processed data to cutting-edge research articles in newly developing fields. This information has been assembled as a matter of public responsibility by the individuals and institutions of the scientific and engineering communities, largely with the support of public funding.

Data are the building blocks of knowledge and the seeds of discovery. They challenge us to develop new concepts, theories, and models to make sense of the patterns we see in them. They provide the quantitative basis for testing and confirming theories and for translating new discoveries into useful applications for the benefit of society. They also are the foundation of sensible public policy in our democracy. The assembled record of scientific data and resulting information is both a history of events in the natural world and a record of human accomplishment.

The recent advent of digital technologies for collecting, processing, storing, and transmitting data has led to an exponential increase in the size and number of databases created and used. A hallmark trait of modern research is to obtain and use dozens or even hundreds of databases, extracting and merging portions of each to create new databases and new sources for knowledge and innovation. However, not

only researchers and educators, but all citizens with access to computers and networks, constantly create new databases and information products for both commercial and noncommercial applications by extracting and recombining data and information from multiple sources. The rapid and continuous synthesis of disparate data by all segments of our society is one of the defining characteristics of the information age. The ability of individuals and organizations to use information in a wide variety of innovative ways is also a measure of success of the original data-collection efforts.

Progress in the creation and use of new knowledge for the national good depends both on the full and open availability of government and government funded data, and on fair and equitable availability of data from the private sector. By Ïfull and openÓ we mean that data and information derived from publicly funded research are made available with as few restrictions as possible, on a nondiscriminatory basis, for no more than the cost of reproduction and dissemination. Fair and equitable availability of data from the private sector means that if commercial content providers receive enhanced protections in their databases, that preferential terms of access to and use of those data by researchers, educators, libraries, and other public-interest entities, firmly rooted in our Constitution and legal tradition, are retained and, when necessary, adapted to the digital and online environment. Moeover, legal rules must ensure that private firms cannot by contract and market power override the traditional Constitutional rights of access and use by the research and educational communities.

Why the United States Should Not Be Compelled to Follow a Flawed European Model for Database Protection

It is our view that any domestic legislation, such as H.R. 354 (or its predecessor drafts in this Committee), that seeks to impose the same or "equivalent" legislation to the E.U. Directive on Databases would be unacceptable. The *sui generis* Database Directive, adopted by the European Communities in March 1996, is an inappropriate model for the United States because of the following major problems:

- The creation of an unprecedented, absolute exclusive property right in the contents of databases, which would decrease even public-interest access to data and reduce competition;
- An overbroad definition of databases that potentially includes every information product that has heretofore been freely available from the public domain;
- The use of other undefined terms and concepts, creating significant uncertainties in the lawÌs scope and application;
- The introduction of long and potentially perpetual terms of protection, with a resulting possibility of no evolving public domain from which previously compiled data could be freely used;
- The absence of sufficient public-interest exceptions for the preservation of public-good activities such as research, education, and libraries, as well as significant curtailment of other usersÌ rights;
- No mandatory legal licenses or other limitations requiring sole-source providers to make data available on reasonable terms and conditions, with due regard for the preservation of competition and the public interests of research and education; and
- The introduction of strong civil (and possibly even criminal) penalties for infringement that likely

would have a chilling effect on the full and open exchange of data for research and educational purposes.

It is important to emphasize that these unwarranted restrictions have been placed on access to and subsequent uses of *factual data*, which traditionally have been in the public domain and, for good reason, have not been covered by copyright or other exclusive rights. Moreover, these restrictions apply as well to collections of "works of authorship" such as journals, textbooks, and anthologies, thus superseding copyright protection. In the case of the research and educational communities, the potential negative effects are exacerbated by the fact that most sources of scientific data are natural monopolies, either because the data contents are unique and not reproducible, as in the case of all observational data of transitory natural phenomena, or they are generated for esoteric niche markets that have a customer base too small to support more than one producer or supplier.

Our concerns are further amplified by the fact that the sui generis restrictions apply as well to publicly funded data in Europe and that this could lead to tremendous strains, or even the breakdown, in certain areas of scientific cooperation between the United States and Europe. Such cooperation is becoming increasingly important for accelerating scientific progress and for sharing costs in such areas as genomic research and global remote sensing studies, yet signs of this tension are already appearing in these important areas of research.

It is possible that the E.U. Directiveìs reciprocity clause will be found to violate the terms of existing intellectual property and trade conventions regarding national treatment of law requirements. However, even if a legal challenge to the reciprocity provision were to fail, other countries, especially those in the developing world, may begin to institute their own sui generis intellectual property rights without national treatment, and discriminate against foreign innovators. Such a result could quickly undermine the now universal norm of national treatment, which was a principal goal of the recent TRIPS agreement under GATT. Thus, the mere fact that the E.U. has adopted a flawed new legal regime for database protection and coupled it with an unwise, and possibly illegal, reciprocity requirement should not induce the United States to emulate it. Rather, our government should challenge the reciprocity provision and independently craft legislation that targets database piracy in a manner that reasonably balances all legitimate interests.

Legislative Alternatives to H.R. 354

We are pleased that the process of deliberating major changes to the U.S. intellectual property law for databases has become more open and appears to have slowed to a rational pace, that the E.U.'s *sui generis* model is no longer the sole option under consideration, and that the participation by representatives of major potentially affected end-user groups, as well as by a broader cross-section of the commercial database and information services industry, has become institutionalized. We especially wish to draw your attention to an alternative legislative approach, "The Database Fair Competition and Research Promotion Act of 1999," which was introduced into the Congressional Record by Senator Orrin Hatch, Chairman of the Senate Committee on the Judiciary, on January 19, 1999, and which we support, as well as to the outcome of Senate Judiciary Committee negotiations on database protection legislation last summer, in which some key concerns of the scientific and educational communities were addressed.

Before discussing these two important developments, we wish to note that we were encouraged by the

two changes that already have been made to this Committeeìs previous version of this legislation, H.R. 2652 (and subsequently Title V of H.R. 2281), in H.R. 354: (1) trying to eliminate the potential for indefinitely prolonging the 15-year duration of protection in section 1408 (c), and (2) expanding the scope of the exemption for certain nonprofit educational, scientific, or research uses in section 1403 (a). The first revision addresses one of the Constitutional defects that was pointed out by various critics of last yearìs version of this bill; the second one responds to some of the concerns we had conveyed last year regarding the potential negative impacts of the legislation on public interest uses, generally, and on our nationìs research activities, specifically. Despite these positive developments, we remain troubled by the scope and substance of a number of the provisions in H.R. 354, and by its approach to addressing the problem of database piracy overall, which seeks to maintain legal equivalency to the E.U. Database Directive.

As you know, in late July of last year, Senator Hatch invited representatives of the various stakeholder groups to participate in a series of closed negotiations, which lasted from the beginning of August until early October. This process resulted in a series of legislative drafts, culminating in a version dated October 5, 1998, which was introduced into the Congressional Record by Senator Hatch on January 19, 1999 as well (referred to as the "Hatch Database Draft" below). Because of the importance of this legislation to the interests of the research community, the Academies took the unusual step of participating directly in these negotiations. We submitted a series of analytical commentaries and specific alternative proposals (see Appendixes A1, A2, and A3 for an abridged selection of those submissions), almost all of which remain relevant to H.R. 354.

The other concerned parties to the negotiations, including a broad cross-section of nonprofit and industry organizations and companies, also submitted constructive proposals in good faith and these were given due consideration by the Senate Judiciary Committee. Perhaps most significant, the Administration provided a consensus critique of H.R. 2652 in an August 5, 1998 letter from the Department of Commerce General Counsel, Andrew Pincus, to Senator Hatch. In addition, the Department of Justice submitted a legal memorandum to Senator Hatch on July 28, 1998 regarding the Constitutional problems of the legislation, and the Chairman of the Federal Trade Commission, Robert Pitofsky, wrote to the Honorable Tom Bliley, Chairman of the House Committee on Commerce on September 28, 1998 concerning its the potential anticompetitive effects.

Although the direct negotiations of the stakeholder parties produced no major breakthroughs or compromise solutions, the final phases of the negotiations, as mediated by the Senate staff, resulted in the October 5, 1998 draft, which produced some far-reaching modifications to Title V of H.R. 2281. Most of these changes substantially implemented important aspects of the Academies' own position, the highlights of which may be summarized as follows:

- The quasi exclusive property right approach of H.R. 2281 was ultimately abandoned in favor of a more reasonable "misappropriation" (unfair competition) approach (see Appendix B, a letter from Professor Harvey Perlman to Senator Hatch dated September 4, 1998, for a critique of why Title V of H.R. 2281--and the current H.R. 354--is not an unfair competition law). This was accomplished by conditioning liability on acts that "cause *substantial* harm to the actual or neighboring market" of database proprietors (section 1302 of the Hatch Database Draft, emphasis added) and by inviting courts, in the legislative history, to determine "substantial harm" in light of "whether the harm is such as to significantly diminish the incentive to invest in gathering, organizing, or maintaining the database" (see the proposed Conference Report Language on section 1302 in the Hatch Database Draft).

- A full carve-out that would immunize customary scientific and educational activities was adopted in place of the weak exception provided under section 1303(d) of H.R. 2281, and the limited and unacceptable "fair use" approach that the Administration had recommended during the Senate negotiations. We considered a "fair use" approach, modeled on copyright law, as inadequate because other basic copyright immunities and exceptions, including the idea-expression dichotomy, are not carried over into the database protection environment. On the contrary, because the proposed database law would protect collections of facts and data that are ineligible for protection under our copyright law, most scientific activities that were previously permissible would become infringing acts under such a law. The burden would then be on scientists to show that a vague fair-use exception should excuse some of these infringing acts from whatever test of harm was adopted. In contrast, we successfully argued that traditional scientific activities should remain unhampered by any new database protection law, as the Administration's consensus position in the August 5 *Pincus letter* also maintained. To this end, section 1304 of the final version of the Hatch Discussion Draft stated that **"nothing in this chapter shall prohibit or otherwise restrict the extraction or use of a database protected under this chapter for the following purposes:**

  1. **for illustration, explanation, or example, comment or criticism, internal verification, or scientific or statistical analysis of the portion used or extracted; and**
  2. **in the case of nonprofit scientific, educational, or research activities by nonprofit organizations, for similar customary or transformative purposes"** [emphasis added].

  Only if a scientist were to cause substantial harm to the database maker by using unreasonable and non-customary amounts of the collection for a given purpose, or if the scientist in fact made a substitute for the original, or otherwise sought to avoid paying for the use of research tools devised as such, would liability kick in. Under this approach, the burden would be on publishers to show that scientists had crossed the line of permitted, traditional, or customary uses, which are immunized. Our operating principle that science should be left no worse off after enactment than it was before, would thus have been substantially implemented. This same line of reasoning extends to our preference for this language over that proposed in section 1403(a) of H.R. 354.

- Additional immunities and exceptions favoring certain instructional and library uses of databases also were defined (see section 1307 of the Hatch Database Draft), although we believe that greater flexibility would need to be given educational users in this context.

- The need for regulation of licensing terms and conditions was expressly recognized in a series of provisions requiring periodic studies of the misuse doctrine (see Sec. 4 and Proposed Conference Report Language, pages 36-37, in the Hatch Database Draft). It is our view, however, that these restraints on licensing should have been codified in the operative clauses of the Act itself.

- The legislative history also clarified the definition of databases in ways that tended to exclude ordinary literary works, and it denied protection "to any ideas, facts, procedure, process, system, method of operation, concept, principle, or discovery, as distinct from the collection that is the product of investment protected by this Act" (see page 31 in the Hatch Database Draft). Again, in our view, it would be much better to codify this definition expressly in the Act itself.

We considered these revisions to Title V of H.R. 2281, while not necessarily optimal, to be

representative of the progress that could be made in achieving a more balanced database antipiracy legislation. Nevertheless, there were other important provisions of the legislation that still required substantial work to make H.R. 2281, and its successor bill, H.R. 354, even marginally acceptable, including, among other:

- the blanket prohibitions on traditionally legitimate commercial value-adding uses;
- the retroactive application of the legislation;
- the incomplete government data exemption, particularly for government databases disseminated by the private sector;
- the excessive length of the term of protection in light of the breadth and depth of the scope of protection;
- the absence of any reasonable limitations on the greatly increased market power granted by this legislation to sole-source data providers; and
- the lack of adequate definitions regarding important terms.

Although both the extent of progress in the Senate, as well as the unresolved issues, indicate that a great deal more work would need to be done on H.R. 354 to bring it into some reasonable balance among all the legitimate competing interests, we believe a better alternative, as noted above, was introduced into the Congressional Record by Senator Hatch. Without going into extensive detail at this time about the relative merits of the two approaches, we wish to emphasize that we consider the approach taken in "The Database Fair Competition and Research Promotion Act of 1999" to be preferable because it:

- Targets database piracy by using true unfair competition principles, without creating unprecedented new property rights in data and unwarranted control in downstream uses of data;
- Maintains a reasonable balance between the interests of database producers and users, including legitimate and economically important value-adding activities;
- Preserves essential public interest uses, including customary scientific, educational, and library activities;
- Adheres to all Constitutional principles; and
- Provides protection against monopolistic pricing by sole-source data vendors in situations where competition is not a de facto reasonable method of sustaining balance of economic interests.

We trust that you and your Committee will review this alternative carefully to make your own determinations. We believe it is especially worthy of note that this alternative is supported by an impressive array of not only research, educational, library, and consumer organizations and institutions, but by many commercial publishers and information service providers.

Later this spring, the National Research Council will publish two reports that will address in greater depth many of the fundamental issues regarding intellectual property rights in the networked

environment, and reviewing the policy options for promoting access to and use of scientific and technical data for the public interest. Also toward the end of this year, AAAS will issue recommendations of an expert group it has convened on the connection between intellectual property and electronic publishing in science.

We hope that these studies will help promote a deeper understanding of the issues underlying the current debate, and we look forward to continuing to work with Congress in this important area. Thank you again for providing us with the opportunity to testify at this hearing.

BIOGRAPHICAL SUMMARY OF JOSHUA LEDERBERG

Joshua Lederberg, a research geneticist, is Sackler Foundation Scholar and President-emeritus of Rockefeller University in New York. Dr. Lederberg attended Columbia P&S Medical School, and he received his Ph.D. in microbiology at Yale. He served as professor of genetics at the University of Wisconsin, and then at Stanford School of Medicine, before coming to the Rockefeller University in 1978. His life-long research, for which he received the Nobel Prize in 1958 (at the age of 33), has been in genetic structure and function in microorganisms. He has been actively involved in artificial intelligence research (in computer science) and in the NASA experimental programs seeking life on Mars. He has also been a consultant on health-related matters for government and the international community, e.g., having had long service on WHO's Advisory Health Research Council. He has been a member of the National Academy of Sciences since 1957, was a charter member of the Institute of Medicine, and was elected Fellow of the American Association for the Advancement of Science in 1982. He has served as Chairman of the President's Cancer Panel, and of the Congress' Technology Assessment Advisory Council, as well as on numerous other consultative panels, including the regents of the National Library of Medicine. He received the U.S. National Medal of Science in 1989. From 1978 to 1990, Dr. Lederberg served as president of the Rockefeller University, where he continues his research activities in the genetics and evolution of bacteria and viruses.

- ARL Federal Relations and Information Policy
- Copyright and Intellectual Property Table of Contents


ARL Home