POLICY FORUM: INTELLECTUAL PROPERTY

Database Protection and Access to Information

William Gardner and Joseph Rosenbaum

Science thrives on open access to compilations of data in electronic or other databases. Thus, scientists should be concerned about the Collections of Information Antipiracy Act recently passed by the U.S. House of Representatives (*1*). If enacted into law, the act would significantly increase the property rights of database owners, which could limit scientists' access to data. Here we discuss the implications of this legislation for scientists and the public. At stake is the appropriate compromise between two competing public goods (*2*). On the one hand, laws that protect data compilations may protect legitimate interests of database providers. Such laws benefit the public by increasing the supply of good data, assuming that the providers deliver data to the public at a reasonable price. On the other hand, the advancement of science, scholarship, and informed public discourse require open access to facts.

The Antipiracy Act

This legislation would amend federal law so that a person who extracts a substantial part of a database for an unauthorized commercial use would face civil liability for damages to the owner of the database and in some circumstances would face criminal charges. To be liable, the unauthorized user must harm the database owner's actual or potential market for a product or service that incorporates that collection of information. To qualify for protection, the database producer must have invested substantial resources or efforts to gather, organize, or maintain the collection. This protection would last 15 years "after the investment of resources that qualified the portion of the collection of information for protection." The bill exempts not-for-profit educational, scientific, or research users of the database from liability but only when those users do not harm the actual or potential market for a good or service. Other exemptions from liability include the use of "an individual item of information, or other insubstantial part of a collection," the independent gathering or use of information, the use of information to verify the accuracy of information gathered, and news reporting. The law exempts news reporters from liability except when they directly compete with the database owner by providing time-sensitive information to an existing market. The law also permits a court to reduce or remit a defendant's payment of damages when (i) the defendant was working for a nonprofit educational, scientific, or research institution and (ii) the defendant believed that the extraction of data in question was permitted under the act. Finally, the act does not cover government collections of information or computer programs for creating or maintaining collections of information (*3*).

Copyright Protection of Databases

Prominent proponents of the legislation include the American Intellectual Property Law Association, the Information Industry Association, the American Association of Publishers, and many commercial database providers and publishers. They argue that current law does not adequately protect databases that do not reflect creative selection or organization of information. They contend that competitors will misappropriate data from such databases, destroying the market for information (*4*).

U.S. law has protected databases from unauthorized copying since the enactment of the first Copyright Act. Courts have given compilations broad protection, requiring competitors to gather the information independently from primary sources. However, the 1976 Copyright Act required that a protected compilation must include original selection, coordination, or arrangement of information. This protection of compilations under this act was most recently addressed by a 1991 decision of the U.S. Supreme Court, *Feist Publications* v. *Rural Telephone Service Co.* [499 U.S. 340 (1991)]. The Court held that a telephone directory (at issue in the *Feist* case) was not copyrightable because it lacked creativity. The Court viewed the directory as an extreme case that fell in a narrow category of works lacking in originality, and it stated that the vast majority of compilations would pass the originality test. However, the Court also argued that only the selection and arrangement of facts in a compilation were protected by copyright, not the facts themselves.

The Antipiracy Act reverses course by allowing database owners to sue persons who misappropriate data from a compilation. The act would protect many databases not currently protected by copyright, because it requires only that the data be "gathered, organized, or maintained by another person through the investment of substantial monetary or other resources." Advocates have argued that these additional protections make it easier to prosecute database pirates whose activities are being facilitated by advances in information technology.

Advocates are particularly concerned that proposed database protection laws in Europe will protect only databases produced in the European Union. Thus, an unscrupulous European producer could acquire an uncopyrightable U.S. database and reproduce and market it from Europe. The U.S. database owner would have no legal means to stop this misappropriation. Advocates believe, however, that if the Antipiracy Act became law, the Europeans might agree to extend their database laws to protect uncopyrightable U.S. databases from infringement in Europe.

The act would also preempt state laws on database protection and standardize rights in the use and ownership of data collections across the United States. Finally, advocates believe that unauthorized use would be policed and prosecuted more consistently and effectively under a federal statute. The act's supporters argue that its increased protection will attract increased investment in databases.

Risks to Science

Critics of the law include the AAAS, the National Academy of Sciences, the Association of Research Librarians, the Digital Future Coalition, the Electronic Freedom Foundation, and many law professors. We and other opponents of the act believe that the law should protect database owners from misappropriation of data by competitors. Our concern is that the extension of copyright protection to compilations that do not reflect original or creative work would lead to inappropriate and intrusive limits on the flow of scientific information or raise the costs of using such information to a level that would harm scientific progress.

The conditions that trigger the protection of the statute broaden the scope of property rights in information in a manner that would obstruct scientists' access to information. Unlike copyright law, the Antipiracy Act contains no exemption for

W. Gardner is in the Departments of Medicine and Psychiatry of the University of Pittsburgh School of Medicine and is the American Association for the Advancement of Science (AAAS) Chair of the National Conference of Lawyers and Scientists (NCLS), a joint committee of the AAAS and the American Bar Association. E-mail: gardnerwp@msx.upmc.edu. J. Rosenbaum is an attorney with Pryor Cashman Sherman & Flynn, LLP, and is a member of the NCLS.

"fair use" of information. Although the bill exempts scientific uses of a collection, it does not exempt such use if it harms an actual or potential market. Consider the hypothetical case described in the box. Dr. Jones made no commercial use of AGD's data. Moreover, AGD had not required Dr. Jones to agree to a contract or license restricting her circulation of the data when she purchased them. Nevertheless, Dr. Jones became liable under the act when she sent copies of the data to other scientists as part of her software application, because by doing so she allegedly harmed AGD's actual or potential market for those data. The current lack of definition of what constitutes a potential market may expose currently legitimate information collection activity to liability.

In theory, Dr. Jones could have independently collected the sequence data from the primary sources and avoided liability under the act. However, Dr. Jones' NIH grant did not include funding for such activities. Moreover, no agency would provide funds to replicate an existing compilation when they could otherwise support original research. The hypothetical court found that Dr. Jones had violated the act by distributing a program that included data purchased from AGD. The judge reduced the damages sought by AGD, as the act permitted when the defendant was a researcher. AGD, however, had never expected to win significant damages from Dr. Jones. Their purpose was to deter scientists from circulating copies of the AGD data, and in this they succeeded. AGD punished Dr. Jones severely by simply filing suit. She ended the trial in debt to her lawyer, and every hour spent on her trial came out of the time remaining on her tenure clock.

One cannot strengthen the rights of information owners without reducing those of scientific users and vice versa. We believe, however, that current law properly balances the rights of information providers and information users. It protects ownership rights in tangible records of facts while promoting the free circulation of the facts themselves. This balance derives from the Constitution, which authorizes Congress to pass laws to "promote the Progress of Science and useful Arts by securing for limited times to authors and inventors the exclusive rights to their respective writing and discoveries." The Antipiracy Act itself acknowledges an important limitation on property rights of database owners by exempting news reporting that uses information in a database, even if such reports damage an actual or potential market. At the very least, the act should protect scientists as extensively as it protects journalists. For example, it should be written to permit scientists like Dr. Jones to make copies as part of the ordinary practice of research, so long as the scientists are not engaged in commercial competition with the database owner. The problem in the vignette is that although AGD is neither an author nor an inventor, the act would give it an effective monopoly on the use of information that was critical to a branch of science.

Advanced Genetic Data (AGD) has compiled data on variation in human DNA sequences and sells access to these data to pharmaceutical firms and other biotechnical customers. The company made a considerable investment to compile the database from their research and from publicly available databases.

Dr. Susan Jones is a molecular geneticist funded by NIH. She has developed a software application that detects whether DNA samples contain members of a library of biologically significant target sequences. The sequence library is stored in a database that is a component of Dr. Jones's application. Dr. Jones compiled her library from various sources, including sequences purchased from the AGD database. After the publication of her work, she shared her application, including the sequence library, with colleagues working on similar problems.

At about the same time, AGD tripled the price of accessing the data set. AGD also filed suit against Dr. Jones, stating that by sharing AGD's sequence with colleagues, she has harmed their market for the data themselves and the software application embodying the data that AGD had planned to develop.

Blocking scientific progress? A hypothetical case illustrating the pitfalls of the Antipiracy Act.

The Antipiracy Act shifts the constitutional balance in three ways. First, the Constitution requires originality for protection, whereas the act would extend property rights to compilations that lack it. Second, the Constitution grants protection for only limited times. The act attempts to meet this requirement by limiting protection to 15 years after the investment of resources that qualifies a portion of a collection for protection. However, because the act protects a discrete portion of a database from the date of investment in that portion, continuously updated databases will be mixtures of protected and unprotected data. Thus the act gives effectively "rolling" and perpetual protection to a collection, which is directly against the Constitution's principle of a limited period. If so, compilations lacking originality would arguably have stronger legal shields than patents or works of creative authorship. Finally, the act would shift the goal of copyright law from the promotion of science and creative authorship to the protection of investment.

The information market, however, changes quickly, and the future could well confirm that databases that do not meet the originality requirement of copyright law do need additional legal protection. Even so, we would oppose legislation similar to the Antipiracy Act because of its potential consequences for science. Contract, trademark, trade secret, and misappropriation laws already provide significant protection for private rights in databases that lack creativity. For example, AGD could have asked Dr. Jones to sign a contract or accept a license on her use of the data. Such a license could have explicitly limited the terms under which she could circulate copies. More important for AGD, they could have imposed terms forbidding resale or copying on commercial customers who might otherwise reproduce and market the data. For Dr. Jones, however, the crux of the matter was her right to circulate copies of her program to research colleagues for noncommercial use. Scientific communities construct a common view of nature through the replication of findings and constant criticism of others' methods. Thus, the sharing of methodology and data within a research community is a necessity, not a convenience. In Jones' view, AGD's intent to limit the circulation of the data she had bought directly threatened a fundamental scientific practice.

Hence our concern that the Antipiracy Act could harm science by restricting access to data. Empirical evidence of a failure of the information market to produce useful data, should that occur, would certainly warrant reexamination of the protection of databases. However, the current situation does not compel us to radically change the legal foundations of copyright protection, as proposed in the act.

References and Notes

1. The Collections of Information Antipiracy Act (H.R. 2652) was passed by the House of Representatives in May 1998. As of this writing, it had not yet been acted on by the Senate.
2. For a comprehensive and informative summary of the arguments for and against the extension of database protection laws, see the "U.S. Copyright Office Report on Legal Protection for Databases" (August 1997); available at http://lcweb.loc.gov/copyright/more.html#rpt.
3. Written statement of the Information Industry Association before the Subcommittee of Courts and Intellectual Property, House Committee on the Judiciary, 30 October 1997, available at www.infoindustry.org/ppgrc/doclib/grdoc017.htm. See also L. Tyson and E. Sherry, *Statutory Protection for Databases: Economic and Public Policy Issues* (1997); available at http://www.infoindustry.org/ppgrc/doclib/grdoc016.htm.
4. Database owners can also copyright their compilations under the Berne Convention for the Protection of Literary and Artistic Works and the Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPs). The TRIPs agreement requires World Trade Organization member countries to protect compilations with intellectual creativity requirements similar to those of current U.S. copyright law.

5. We thank M. Frankel, A. Fowler, B. Black, L. Loevinger, and S. Poulter.