

POLICY FORUM: INTELLECTUAL PROPERTY

Europe's Database Experiment

Stephen M. Maurer,* P. Bernt Hugenholtz, Harlan J. Onsrud

For most of the 20th century, database protection was like the big mongrel dog that nobody wanted to wake. U.S. and European lawyers agreed that intellectual property law should not apply to ordinary facts. Information could only receive protection if it was "original" or "novel." This kept most facts in the public domain where scientists, inventors, and entrepreneurs could use them.

The U.S. debate started in 1991, after the Supreme Court held that competitors could extract copyrighted information from a Kansas telephone book (1). At first, the database lobby was so weak that Congress basically ignored them. Then, in 1996, the European Community (EC) issued a directive telling its members to create a new type of database right, and threatened to withhold protection from American firms unless the United States followed suit (2). Congress took the EC's threat seriously. Strong, European-style legislation came close to passing in 1999 and again in 2000. Three factors stopped it. First, many members of the database industry (for example, Yahoo) opposed strong legislation because they were afraid it would make data collection harder and more expensive. Second, the policy arguments for protection were murky. In fact, most scholars believed that there were strong empirical (3) and theoretical (4) reasons why a new law could create more problems than it solved. Finally, the House Commerce and Judiciary Committees could not agree on what type of statute was needed. (5). By early 2001, chances for new legislation seemed to be fading (6).

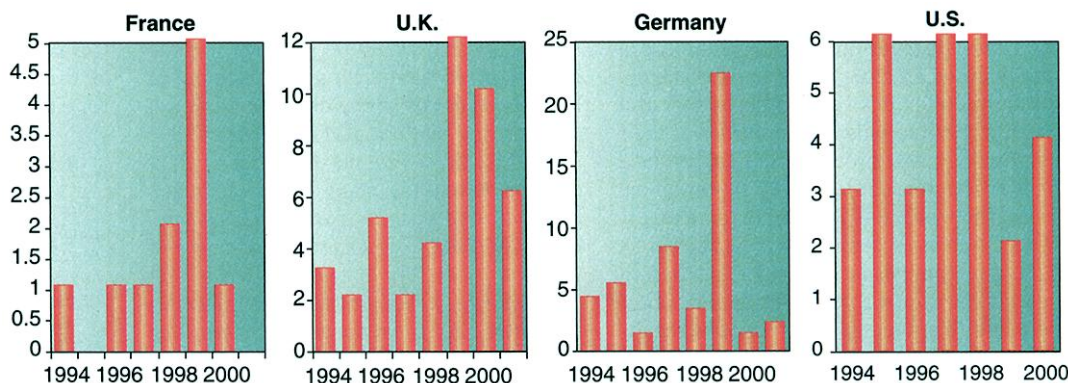
Things changed dramatically on 29 March 2001 when the relevant committee chairmen told a surprised Washington

press corps that they were "committed" to "getting a database protection bill ... passed by Congress this year." (5). Chances for substantial new database protections are higher now than they have been in years (6).

The irony is that Congress still does not know what the EC wants it to imitate. Like most new statutes, the Council Directive is a work in progress. What do we really know about the EC's experiment?

Europe Under the Directive

Like all intellectual property laws, Europe's directive was designed to address a paradox. Normally, society opposes mo-



A European growth spurt? New companies entering four large database markets according to a leading commercial catalog (18). Unlike the United States, all three EC countries show a sharp, one-time growth spurt after host governments implemented the EC Council Directive in 1998. Further details can be found in (10).

nopolies because they create artificial scarcity and raise prices for consumers. Intellectual property, on the other hand, creates monopolies to encourage new products. The trick is to get the best possible bargain by restricting new rights to products that are valuable and cannot be obtained by other means. Careful legislators do this by imposing threshold requirements (such as "novelty" and "creativity") that dole out rights as sparingly as possible.

The EC Council Directive was supposed to have two of these threshold requirements. First, it limited the concept of a "database" to "collections of independent data ... arranged in a systematic ... and individually accessible" way. Recent court decisions show that this definition has become almost irrelevant. Almost any jumble of data, including pages in a Web site and classified ads in a newspaper, will suffice (7). Second, the directive limits

protection to owners who make a "substantial investment" in their data. European courts have acknowledged that this standard is also "fairly low" (8). Even trivial databases, such as a collection of 251 hyperlinks, are routinely protected (7).

These cases did not start the erosion of the public domain, but they have clearly accelerated it. For now, the biggest challenge is in biotechnology. Companies worry about committing hundreds of millions of dollars to a project if there is the slightest chance that the underlying data could belong to someone else. For this reason, some companies warn employees never to acquire data by surfing the Web. Until recently, this kind of rule worked surprisingly well. However, post-genome science is changing. In order to make progress, scientists will have to combine data from dozens of academic and commercial sources. Under these circumstances, discovery could soon be limited to the pace at which lawyers write contracts (9).

One of the most amazing things about Europe's database experiment is that at least 50% of all lawsuits have been brought by the tiny minority of companies (fewer than 5%) that own telephone listings, sporting event dates, concert times, and broadcast schedules (10). What does this hyperactive minority have in common? None of them actually collect data from the outside world. Instead, they make it up. We call this phenomenon "synthetic data." The classic example of synthetic data is a telephone number. Unlike, say, nuclear physics or genome data, would-be competitors cannot obtain this information through independent research. Copying the telephone book is their only option. This has important consequences. Normally, the threat of "independent invention" provides a powerful brake on the prices that owners can charge (11). Under the Council Directive, synthetic data are often more valuable than genuine information.

S. M. Maurer is a practicing attorney at 2632 Hilgard Street, Berkeley, CA 94709, USA. H. Onsrud is in the Department of Spatial Information Studies, University of Maine, Orono, ME 04469-5711 USA. P. B. Hugenholtz is in the Institute for Information Law, University of Amsterdam, Roeter 84, NL-1012 KX Amsterdam, The Netherlands.

*To whom correspondence should be addressed. E-mail: maurer@econ.berkeley.edu

Although scientists rarely use synthetic data, scientific databases have their own independent invention problem. Resources like weather data, geologic maps, and the U.S. census have almost no chance of being funded a second time. Many scientists and organizations, notably the International Atomic Energy Agency, are worried that private publishers could capture and monopolize such data (12).

Many companies dislike databases that help consumers find out about competing products. During the 1990s, online indexes began appearing that allowed consumers to conduct unified searches across multiple competing Web sites. Once a consumer finds an interesting "hit," he or she can go to the underlying Web site by clicking on a link. In principle, these "deep linking" sites are a tremendous tool for unifying data. In practice, database owners have repeatedly sued to break the links in order to keep their online news stories, want ads, and real estate listings separate. Particularly in Germany, most of these suits have succeeded (7, 13).

Data aggregation is also becoming a problem for science. Over the past 5 years, academic (for example, Swiss-Prot) and commercial biologists have burdened their data with "pass through" rights that place restrictions on data even after they have been incorporated into other databases. This creates opportunities for gridlock. A U.S. government data provider has had to restrict its use of Swiss-Prot data in order to avoid the potential repercussions resulting from redistribution.

EC officials say that legislation has helped the European database industry to grow at "enormous rates" (14). As far as we can tell, the evidence comes down to one official's informal visits to Web sites and department stores. He reports that European databases surged in 1998 and have continued to grow "at a more modest rate" since then (15). For now, the EC is keeping further details confidential (15). However, there may be an added twist. According to research by one of us (S.M.M.), European database production returned to predirective levels almost immediately (10). If so, the net effect appears to have been a modest, one-time boost (see the figure on page 789).

A Divided House?

People assume that the modern world is not big enough for two different database regimes. This ignores the fact that almost all commercial data (for example, telephone numbers, credit histories, and Nielsen ratings) are local (6). For a business that collects and sells statutes in Iowa, the EC could just as well be located in another solar system.

Commercial science is the most likely place where EC and U.S. rules could clash. Even here, the fallout would be minimal. U.S. firms depend on a sophisticated mix of strategies to stay ahead of would-be copiers, including contracts, download restrictions, and frequent updates. Europe's legislation does not affect these strategies in any way. Furthermore, the Council Directive contains a loophole: If a U.S. company wants database rights, it can get them by moving some of its operations to Europe or else by finding a local partner. So far, commercial bioinformatics firms, who have the greatest risk, have decided that the extra protection is not worth disrupting their businesses (9, 10).

Fixing the Directive?

In hindsight, the 20th century's genial neglect of database protection issues looks like a pretty good thing. We think that for the EC the best course would be to admit that the directive is unnecessary and repeal it. However, this may not be politically possible. What else can Europe do? Probably the easiest step would be for lawmakers to minimize the database right by passing as many exemptions as the directive allows. It is surprising that not all EC countries have done this. For example, France and Italy have not taken advantage of the directive's optional exemptions for scientists who download data for educational or research purposes (16). The basic problem is that such reforms, although useful, are inherently limited. For most scientists, having the right to download data does not mean much if the extracted information cannot later be reutilized and republished. Creating a useful scientific and educational exemption is impossible under the current directive.

Some Dutch courts are trying to stretch the EC Council Directive in a different way. They refuse to extend protection where the owner would have made its "substantial investment" anyway as part of some other, nondatabase business (7). This "spin-off" rule eliminates protection for synthetic data like telephone listings and sports schedules. Because most other European courts have resisted the idea, a final answer will probably have to come from the European Court of Justice. A second approach would be to require monopolists in synthetic data to sell their information to competitors on fair and nondiscriminatory terms. This idea appeared in early drafts of the Council Directive and could easily be restored. Although establishing a "fair" price in court is notoriously difficult, this solution would also be a big improvement.

Finally, why not limit protection to databases that need it? Judges in the United States and Europe have already devel-

oped an "unfair competition" doctrine which says that courts should only take action if unfettered copying would threaten the product's existence (17). Early drafts of the Council Directive would have adopted a similar principle by limiting protection to "unfair extraction." Turning the directive into an unfair competition statute would also be good politics. North America and Japan are understandably leery of a right that the EC invented out of whole cloth. On the other hand, unfair competition laws have been around for years. If the EC wants a model for global consensus, it need not look any further than its archives.

Congress's Choice

We have discussed how the EC Council Directive may have given Europe's database industry a limited one-time boost. If so, the cost was high. Recent court rulings show that the directive has eroded the public domain, overprotected "synthetic value" of doubtful worth, and raised new barriers to data aggregation. Congress should take a long, hard look at these drawbacks before imitating Europe's database experiment.

References and Notes

1. *Feist Publications, Inc. v. Rural Telephone Service Co., Inc.*, 499 U.S. 340 (1991).
2. Council Directive No. 96/9/EC, O.J.L 77/20 (1996).
3. Y. Benkler, *Berkeley Technol. Law J.* **15**, 435 (2000).
4. S. Maurer, S. Scotchmer, *Science* **284**, 1129 (1999).
5. Transcripts are available at www.techlawjournal.com/cong107/database/20010329.asp
6. J. Reichman, personal communication.
7. P. B. Hugenholtz, "The new database right: Early case law from Europe" (2001); available at www.ivir.nl/publications/hugenholtz/fordham2001.html
8. *British Horseracing Bd. Ltd. v. William Hill Org. Ltd.* (Chancery Div. 2000), available at http://www.courtservice.gov.uk/judgments/judg_frame.htm?OpenDocument;
9. Lee Bendeckey, vice president and general counsel of Incyte Genomics, Inc., personal communication.
10. S. Maurer, in *Intellectual Property and Innovation in the Knowledge-Based Economy*, J. Putnam, Ed., proceedings of the conference 23 to 24 May 2001, Toronto, ON, Canada (Univ. of Calgary Press, Calgary, AB, Canada, in press).
11. S. Maurer, S. Scotchmer, *Olin Working Pap. Ser. Law Econ.* 98-11 (1998).
12. R. B. Firestone, staff scientist, Ernest Orlando Lawrence National Laboratory, personal communication.
13. Collected German court opinions are on file with one of the authors (P.B.H.); see also, T. Heyden, *Copyright World* **102**, 7 (2000).
14. J. Reinbothe, untitled paper presented at WIPO's "Protection of Databases Workshop" on 16 September 1999, Geneva; available at <http://ecommerce.wipo.int/meetings/1999/papers/reinbothe.html>.
15. J. Gaster, personal communication.
16. P. B. Hugenholtz, in *Intellectual Property and Information Law: Essays in Honour of Herman Cohen Jehoram*, J. J. C. Kabel and G. J. H. M. Mom, Eds. (Kluwer Law International, The Hague, 1998), pp. 183-200.
17. *National Basketball Assn. v. Motorola, Inc.*, 105 F.3d 841 (2d Cir. 1997).
18. *Gale Directory of Databases* (Gale Group, Farmington Hills, MI, 1993 to 2001).