Campus Database Issues

Monograph 99-6

Robert O'Neil
University of Virginia
School of Law
Thomas Jefferson Center for the Protection of Free Expression
580 Massie Rd.
Charlottesville, VA 22903
804-295-4784

$5.00 © Robert O'Neil, 1999
University of Houston Law Center/Institute for Higher Education Law and Governance (IHELG)

The University of Houston Institute for Higher Education Law and Governance (IHELG) provides a unique service to colleges and universities worldwide. It has as its primary aim providing information and publications to colleges and universities related to the field of higher education law, and also has a broader mission to be a focal point for discussion and thoughtful analysis of higher education legal issues. IHELG provides information, research, and analysis for those involved in managing the higher education enterprise internationally through publications, conferences, and the maintenance of a database of individuals and institutions. IHELG is especially concerned with creating dialogue and cooperation among academic institutions in the United States, and also has interests in higher education in industrialized nations and those in the developing countries of the Third World.

The UHLC/IHELG works in a series of concentric circles. At the core of the enterprise is the analytic study of postsecondary institutions—with special emphasis on the legal issues that affect colleges and universities. The next ring of the circle is made up of affiliated scholars whose research is in law and higher education as a field of study. Many scholars from all over the world have either spent time in residence, or have participated in Institute activities. Finally, many others from governmental agencies and legislative staff concerned with higher education participate in the activities of the Center. All IHELG monographs are available to a wide audience, at low cost.

Programs and Resources

IHELG has as its purpose the stimulation of an international consciousness among higher education institutions concerning issues of higher education law and the provision of documentation and analysis relating to higher education development. The following activities form the core of the Institute’s activities:

Higher Education Law Library

Houston Roundtable on Higher Education Law

Houston Roundtable on Higher Education Finance

Publication series

Study opportunities

Conferences

Bibliographical and document service

Networking and commentary

Research projects funded internally or externally
Campus Database Issues

by Robert M. O’Neil

Though this topic holds keen academic interest for me, I must confess my contact with it has been anything but academic. Several years ago, right after leaving my last administrative post, I agreed to serve on an academic senate committee to revise our university’s faculty handbook. When we reached the section on library digital and electronic resources, I suddenly found myself the target of greater faculty hostility than any of my presidential edicts ever provoked.

How, asked this random group of mainly Arts & Sciences professors, could I defend a library policy that provided unlimited free access to LEXIS/NEXIS and Westlaw only for me and my law school colleagues? The short answer was easy -- those were the only terms on which West and Mead (now Elsevier) would make access available to a large and complex institution. It was either the law faculty only or no faculty at all. My handbook committee colleagues gradually turned to other issues, though throughout our deliberations they remained deeply resentful of someone’s unconscionably restrictive database access policies. That was my initial exposure to the real-world dimensions of this issue, and a clear reminder that controversy about it has not been confined to the political realm.
The issue has become steadily more complex since that encounter. Everyone knows how dramatically the contents of electronic databases have grown during the past decade. Between 1991 and 1997, for example, the number of files in such databases increased from some 4 billion to over 11 billion, nearly a two hundred percent growth. The number of searches of such information centers grew comparably -- roughly doubling during any five year period during the decade\(^1\). What is less familiar is how dramatically the source of such databases has shifted. It used to be that universities were the creators; in the late 1970s 78% of all databases were produced by government, academic and other non-profit providers, with the commercial sector accounting for the remaining 22%.

\(^1\) See Database Coalition, H.R. 1858 Data, at http://www.databasecoalition.org/hr1858/legalprt/legalprt.html, pp. 3-4.\ldots
By the end of the decade, that balance had altered dramatically. By 1997 (the last year for which we have detailed figures) the numbers were exactly reversed; the government/academic/nonprofit sector now accounts for barely twenty percent of database production, while the proprietary world of commercial providers has assumed the dominant role in this rapidly changing environment.\textsuperscript{2} Tensions between erstwhile producers (now mainly users) and the new producers. That tension and its legal dimensions are the focus of this paper.

At the outset, two modest disclaimers seem to be in order. The first relates to subject-matter competence. Though I have at various times taught intellectual property, I gave up that fascinating field four years ago to offer a new course called Free Expression in Cyberspace. I try to keep up with developments in copyright, though with growing difficulty. The other disclaimer is subtler, but probably more critical: Though I have no direct (nor even, so far as I am aware, indirect) financial stake in the fortunes of any database protagonist, my lifelong involvement in the academic community inclines me to a rather partisan view on these issues. Thus the position that seems to me not only congenial, but eminently logical, reflects my experience and attendant bias.

\textsuperscript{2} Id.
In fairness, some people have seen this issue coming for a good while. The 1978 report of the CONTU Commission contained the cryptic comment that “unauthorized taking of substantial segments of a copyrighted data base should be considered infringing,” but left the implications to others. Almost a decade later, the Office of Technology Assessment’s Intellectual Property Rights paper observed that “the unique characteristics of computer database technology may severely curtail the usefulness of copyright protection for works that are distributed on-line, since copyright protection extends only to the work’s expression.” Such early premonitions as these were unlikely to advance the analysis markedly. Nor did they create within the academic community a level of consciousness about impending tensions and conflicts that would, in retrospect, have been most appropriate and helpful.

There are two distinct reasons why the database issue has now attracted the attention of the United States Congress. In the three years since the European Union adopted its database directive, mandating copyright protection within the fifteen member nations, there has been a strong economic incentive for this country

---


to adopt comparable legislation that would enhance access to European markets.\textsuperscript{5} The other catalyst is purely domestic; in the currently uncertain state of the law -- of which more in a moment -- both producers and users seek better and clearer statutory protection of their respective interests in ways that only Congress can provide.

Pending in Congress during the session that ended at the close of 1999 were two bills that would approach the protection of digital databases in starkly different ways. H.R. 354, which mainly reflects the interests of database producers, was introduced last January by North Carolina Representative Howard Coble. This bill would apply retroactively to compilations created as long as fifteen years earlier, and (with certain specific exceptions) would broadly protect “information that has been collected and has been organized for the purpose of bringing discrete items of information together in one place or through one source so that users may access them.”

A major concern of the database user community is that H.R. 354 would effectively preclude, or at least would severely inhibit, transformative uses of data contained in a protected compilation. In the words of a very recent New York Times editorial: “The bill would prohibit most users of a database from making available information from that collection of data if the action would cause financial harm to the original database publisher. This could deter innovators from building new and better databases by recombining, culling or adding information to existing

---

A sharply contrasting approach was embodied in H.R. 1858, introduced last spring by House Commerce Committee Chairman Tom Bliley. In mid summer it received voice vote approval from the House Telecommunications Subcommittee, but went no further. This bill would also penalize piracy of databases, but would do so by targeting specific abuses rather than conferring broad and blanket protection. It would also not prevent transformative uses of existing collections for the purpose of advancing knowledge. Unlike its protagonist, H.R. 1858 would apply only to collections created after its effective date, thus reflecting the constitutional goal of encouraging future innovation. The Bliley bill would rely on civil remedies and FTC enforcement, while the Coble bill contains criminal sanctions as well.

---

Though a full account of the differences between the two bills requires many pages, this brief summary contrasts the bare essentials sufficiently to explain why the producer community supports H.R. 354, while users champion H.R. 1858. Just as this session of Congress was nearing its end, a recent New York Times editorial --- entitled “Fair Use for Databases” --- noted the divergence of philosophy between the two bills, adding that “the issue is increasingly important because electronic databases have become integral to all kinds of commercial and educational activities on the Internet.” The editorial recognized that the distance between the two approaches is enough to make quick or easily solutions unlikely. Both bills, or at least close relatives, are certain to be back in the next session.

The legal antecedents of the current debate now deserve brief review. Even someone who is several years away from teaching copyright knows the basics. Facts as such are, of course, not legally protectible. Compilations of facts may, however, claim limited protection. The eternally elusive question is under what conditions, and to what degree, such a compilation or arrangement of facts may

---

8 See Database Coalition, Side-by-Side Comparison of Database Protection Bills, at http://www.databasecoalition.org/dbside-by-side, containing a twelve-page, single-spaced, item-by-item comparison of the two then pending bills.


claim copyright. That issue received contrasting disposition in the lower courts, though with a mounting consistency in recent years. Despite some nostalgia on the part of database producers, decisions at least since the major copyright law revision of 1976 showed markedly little sympathy for “sweat of the brow” -- the notion that energy, industry and diligence alone might transform a compilation of public domain facts into a legally protectible work.

Eight years ago, in the *Feist* case, the Supreme Court finally entered the fray and offered some guidance. An “original selection or arrangement of facts” might claim copyright, though it would be “limited to the particular selection or arrangement” and in no event protecting any of the facts themselves.11 Gone at last was any notion that “sweat of the brow” invested in discovering or identifying facts might justify copyright. And even where a compilation might exist, its scope would be “thin” at best.

While *Feist* resolved some questions, it raised others and left many still open.12

---

11 Id at .

12 See, e.g., Reichman & Samuelson, Intellectual Property Rights in Data?, 50 Vand. L.
As is the nature of copyright, lower federal courts must still apply broad precepts to infinitely variable fact patterns between the obviously protectible arrangement and the manifestly public domain material. Not surprisingly, the user and producer communities differ on how adequately protective the post-*Feist* litigation has been. There have been a substantial number of relevant cases during the decade and, while headcounts are not helpful, it would be fair to say that database protection has fared reasonably well. Curiously, only one of these recent cases addressed the precise issue that is our concern today and has been the focus of the House bills -- digital or electronic databases. In *Corsearch v. Thomson & Thomson*,¹³ a federal district judge in New York ruled that the compiler of an electronic database of trademark information had “offered sufficient evidence of its selection, coordination, enhancement and programming of the state trademark data, as well as other contributions that establish the originality and requisite creativity, and thus copyrightability, of the . . . database.” ¹⁴

On the basis of this judgment, and many others since *Feist* that protect certain elements of non-digital collections, one may at least question two premises of the industry perspective -- that in *Feist* the Supreme Court reversed a strongly

protective trend in the lower courts, and that the impact of *Feist* has been uniformly hostile to database protection. This is not to say that factual compilations may not need firmer legal safeguards, but only that the courts during the past decade have not been uniformly unsympathetic.

\[14\] Id. at 322.
The central premise of the case advanced by the database industry is that the law ought to provide greater protection for the major investment of time, resources and ingenuity in creating electronic databases -- essentially, that nothing less than H.R. 354's highly protective approach would suffice.\textsuperscript{15} It is thus appropriate to assess the adequacy of alternative safeguards for database creators. Beyond the current and not wholly uncongenial state of copyright law, the user community cites several such alternatives. They are really of two types -- those that might be termed "self help" and those that, in contrast, rely upon legal safeguards and sanctions.

Marshall Leaffer, in the brand new edition of his copyright treatise, notes that "for some time,. compilers of databases have looked to protection for their collections of information outside of copyright law."\textsuperscript{16} Chief among the "self-help" protections seem to be agreements negotiated with -- or terms imposed upon -- users and distributors. A recent article in \textit{IP Magazine} notes that contracts offer potential solace; "many databases, particularly online databases, are distributed subject to licensing agreements under which the licensee -- the user -- agrees not to

\textsuperscript{15} See statements at http://www.infoindustry.org.

redissemate the information. Cross-licensing in this context has obvious and mutual advantages.

The other "self help" safeguards are technological in nature -- encryption, serial copy controls, and watermarking among others. To the extent that the Digital Millennium Copyright Act forbids the circumvention of technological protections, and bans the manufacture of circumvention devices, it offers important protection for all copyrightable databases. And, as the Database Coalition’s current position paper notes, “the ban on circumvention devices helps the few non-copyrightable databases because it eliminates devices which would circumvent technological protections applied to such databases.”

So much for self-help. The remaining alternatives invoke legal provisions apart from copyright. State common law, for example, uniformly provides remedies for misappropriation, which at least may protect news organizations and similar information sources from competitive misuse of “hot news.” Producers tend to minimize the efficacy of such options, not only because of the “hot news” limitation, and the wide variations among states, but also because of continuing doubts about the broadly preemptive effect of section 301 of the Copyright Act -- at least on remedies that are equivalent to the federally created interests. The House Report that accompanied the 1976 revision seemed to encourage, or at least

---

18 See Database Coalition, “H.R. 1858 Data: Legal Protection Now Available to Databases,” at http://www.databasecoalition.org/hr1858/legalprt/legalprt.html

15
condone, survival of state law misappropriation claims, though the courts have been somewhat less sympathetic to such suits. Even so, a complete assessment of database creators’ options should include some mention of state misappropriation law. (Interestingly, H.R. 354 contained, apparently to the industry’s satisfaction, a federal preemption clause that seemed at least as pervasive as section 301.)

Trademark law may also have some potential utility for producers. Increasingly, certain databases have become dominant in their respective fields of information, to a degree that trademark protection may limit unauthorized use by others. As a recent paper posted by the Database Coalition on its extremely helpful website suggests: “Even if a second publisher copies much of the data from the first publisher, consumers will still buy the first database because of the name brand;” accordingly, “trademark law prevents the second publisher from using the first publisher’s name.”19 Of course many producers would reply that their concern is not so much the trade name as the contents -- or that those few producers with such dominant marks probably least need the sort of protection to which measures like

---

H.R. 354 would provide. Even so, trademark is also part of the total equation.

Finally, the user community notes that -- whether or not existing legal protections are entirely adequate -- the database industry has flourished during the 1990s by every conceivable measure. Apart from the impressive trends in database creation and user access, the steady and massive flow of private capital into this sector suggests its attractiveness to investors who, throughout this boom decade, have faced myriad appealing opportunities. Thus the range of responses from the user community might lead even a neutral observer to the view that major new protections are not needed.

Yet all this may really be missing the point, or at least may risk losing the forest in the trees. The basic issue is, as the *Feist* Court reminded us perhaps excessively, a constitutional one. Facts are the property of everyone, and the personal domain of no one. Permitting any legal protection for compilations of facts creates a risk that access to the components of those compilations may be curtailed. Of course *Feist* was right to recognize -- as federal courts had done for decades -- that there may be situations in which denying protection to truly ingenious arrangements of public domain material would be both unconscionable and unnecessary. Yet *Feist* also stressed that it was not “sweat of the brow” which deserved protection, but only creativity and originality, and that even in the
most compelling case such protection would be “thin.” And the post-Feist decisions -- those that have found infringement as well as those that denied protection to compilations and databases -- have made clear the very limited scope of the creator’s or compiler’s legal claims. Clearly this is an area in which the presumption is against protection, no matter how conscientious or industrious the database producer may have been.20

The case against broader protection has been well stated in a sharp rebuke to H.R. 354 by University of Wisconsin-Madison Library Director Kenneth Frazier. In a recent op-ed piece, Frazier notes that “in many cases the public has already paid for the data before it is compiled and packaged in a database,” and then asks rhetorically, “How could anyone own a collection of facts?” He then adds his own special librarian’s concern about comprehensive protection for electronic databases: “The combined power of computers and the Internet has made it possible to compile huge databases and then to charge everyone for gaining access to the information. The combination of new legislation with the technology will, in effect, allow database owners to control how factual information is used, who uses it and at what cost.” Frazier’s deepest foreboding is for the future of scientific

---

inquiry, an interest of special prominence at UW-Madison: "[T]he issue is of critical importance to scientists who depend on using a wide variety of data sources to advance knowledge. In fact, the advancement of science depends directly on the availability and use of data compiled by other scientists."\(^{21}\)

The mid-November *New York Times* editorial expressed a similar concern. Noting that “officials at the National Academy of Sciences have expressed fears that the bill [H.R. 354] could hamper scientific research,” the *Times* cautioned that proposed statutory exemptions designed to reduce such risks “are not enough to keep full access to the facts open.”

That comment leads to a few concluding comments for the good of the order. First and foremost, we in the academic community must continually remind ourselves that we are no longer, as we once were, a community of database producers and information compilers, but are now overwhelmingly database users and are thus increasingly at the mercy of the proprietary institutions that generate such resources, for access to which in most instances we or our institutions must pay ever steeper fees. Thus we have a strong collective interest not only in following legislation that is designed to maintain an equitable balance between producers and users, but actively to support and encourage such legislation.²³

---


²³ A relevant and appropriate response was the decision of the Association of American Universities to retain special legislative counsel to protect and represent the interests of the
Second, we should assume that in the next session of Congress -- and for that matter in every ensuing session until the issue is finally resolved -- there will be offspring or progeny of H.R. 354 and of H.R. 1858 which will clearly differentiate the respective sets of interests and views on database protection. We should at least be far more familiar with the general approach of such bills than, I will confess, I had been and would have been but for this assignment. This is an issue which is too important for us all to be left entirely to administrators -- a comment which you may appreciate is not meant to evidence distrust or lack of respect, but simply to convey the political reality that governing boards and senior administrators cannot single handedly represent the interests of the information-user community. They need help, and that is where the rest of us ought to be more substantially involved than we have been to date.

Third, it will not do to focus solely on the bill we favor --the son or daughter or niece or nephew of H.R. 1858. Since the alternative approach may well ultimately prevail in Congress, it deserves equally close scrutiny. While the broadly protective provisions would not likely be deleted or even substantially altered, what is malleable are the exceptions and exemptions which would then become the crucial safe harbors in this far less friendly ocean. Though the user community widely faulted the exceptions in H.R. 354 as inadequate to permit
certain transformative and other important uses of databases, much less was said or done by way of designing more protective alternatives. In short, we need not only to make clear why we favor the 1858 approach and disfavor that of 354, but need equally to indicate how the disfavored bill could be improved. In that way, user needs and interests stand to be reasonably protected whichever bill eventually prevails.

Finally, we can enhance understanding of the state of the law, both before and since Feist. There seems to be, as I noted earlier, substantial nostalgia about good old days that never were -- times when, reputedly, compilers and collators could claim copyright protection largely if not entirely on the basis of extensive perspiration. That never was the case, and certainly not within any relevant time period. We also know that database producers have fared reasonably well even since Feist -- albeit not quite as well as they might have hoped -- and that the single reported case dealing with an electronic database\(^\text{24}\) came out in the producer's favor. Yet the producer community may have carried this part of the debate almost by default, not by distortion or misrepresentation of the record, but simply by making it appear that the level of protection needed to be strengthened as a

matter of justice and equity. When the truth is fully understood, fewer tears may be shed for those who have labored long and hard to generate valuable databases and factual compilations, but have not seen their efforts go unrequited.