Database Protection: Its Forms and Their Effects

David Moss
Damon May
Don Totten
Elijah Esquibel
Charistel Ticong

CSEP 590: IT Policy
Professors Ed Lazowska and Stephen Maurer

Contents

Introduction

In recent years, a good deal of concern has arisen regarding the right of entities that own databases to control how those databases are used. Many believe that databases require more protection than existing copyright law provides. Without additional protection, assert proponents of database protection, companies are discouraged from investing in the creation of databases because the fruits of their efforts could be stolen by any number of parties. Several countries have recently adopted or considered legislation designed to protect the rights of database owners.

Database owners and their competitors, however, are not the only parties concerned with this kind of legislation. Other groups rely on access to data from many different sources. Additional protection may make their tasks more difficult or even impossible.

Assuming that some level of protection for databases is desirable, what methods are available? This paper discusses several possible mechanisms.

In the US, there are at least three relevant avenues to consider when determining protection methodology:

1. Legislation offering Copyright-like protection to databases
2. Extension of a body of case law known as the "hot news" cases
3. Retention of the status quo.

We will start by providing some background on the current state of affairs in the US (the applicability of copyright law and the hot news cases), discussing the three options above, and touching on the legislation currently under consideration in the United States.

Next we will examine the state of affairs in Europe. In 1996 the European Union enacted a major Directive to bolster the rights of database owners. The Directive has drawn a great deal of criticism from groups who fear its effects on scientists, libraries, and other groups. We examine the track record of the Directive in EU courts, and the first official efforts to evaluate the Directive's effects, and provide some hope for groups concerned about the Directive's direct or indirect effects on their ability to make use of data.

In the next section, we explore the clinical researchers' views on statutory database protection in the form of the Health Insurance Portability and Accountability Acts' Privacy Rule, HIPAA. Our goal is to reconcile the researchers' generally negative views on the subject with the costs and benefits of HIPAA as he perceives them.

Finally we will discuss the pros and cons of *sui generis* database protection. The main players will be examined, as well as their arguments of for and against *sui generis*. The widespread reaction to the results of the *Feist* case have caused a dilemma – what should be available to the public and what should be protected? Current database cases will show the direction the US is headed with regard to *sui generis* protection of databases.

We make a comparison between all of these different approaches, evaluate their effects to date, and show the merits and pitfalls of each.

Copyright Protection in the US

Lead Author: David Moss

Background

On a policy level, there are at least three potential avenues for Database Protection that merit consideration:

1. The body of existing case law protecting "Hot News" may be extended
2. The existing Copyright regime may be imitated; or
3. The current status quo (i.e. lack of specific rules) may be left unchanged.

In order to analyze these options, a brief discussion of copyright protection (for compilations/databases), preemption by the Copyright Act (the "Act"), and the "hot news" cases is required.

The Copyright Act and Compilations

Section 103 of the Copyright Act offers protection to "compilations and derivative works."[i] A compilation is defined as a work formed "by the collection and assembly of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship."[ii] Clearly, a database (if protected under the Act) is to be considered a compilation. Compilations, in order to be protected, must satisfy (among other things) section 102(b) of the Copyright Act (prohibiting the protection of ideas, concepts, principles, etc.).[iii] The question, with respect to compilations, is, to what extent are they protected? The answer is set forth in the statute as follows:

> The copyright in a compilation…extends only to the material contributed by the author of such work, as distinguished from the preexisting material employed by the work, and does not imply any exclusive right in the preexisting material. The copyright in such work is independent of, and does not affect or enlarge the scope, duration, ownership, or subsistence of, any copyright protection in the preexisting material.[iv]

In 1991, the Supreme Court clarified and interpreted the issue of protection for compilations in a case referred to as "Feist."[v] The basic holding of the Court in Feist is that, in compiling a work (i.e. database), effort alone will not trigger protection under the Act; there must be "some minimal degree of creativity."[vi] Even if a compilation is protected, that protection only extends to the *expression* of the underlying facts (the facts themselves are not protected).[vii]

The Copyright Act and Preemption

The Act expressly preempts state law(s) in certain situations.[viii] Case law has clarified that a state law is preempted by the Act where (1) the work that the state law purports to protect "fall[s] within the ambit of copyright protection," and (2) the state law asserts a right or rights equivalent to the exclusive rights protected by federal copyright law(s).[ix] A plain language reading of the statue and case law indicates that any attempt to protect the underlying facts/data in a database is expressly preempted (and so prohibited) by the Act. The "Hot News" cases mark an exception to that rule.

The "Hot News" Cases

(1) The INS Case: The case giving rise to the concept of protection for hot news is referred to as the INS case.[x] In the INS case, the issue revolved around one news reporting agency, INS, allegedly pirating news broadcasts from a different news reporting agency, AP, and redistributing the material to INS subscribers.[xi] It is worth noting that because of time changes and West Coast/East Coast issues, the INS subscribers were sometimes getting the news earlier then were the AP subscribers. The Court in INS held that although there is no right to the underlying news itself (i.e. facts), in that specific situation INS should be prohibited from redistributing AP's material "until its commercial value as news to [AP] and all of its members has passed away."[xii] The holding was based on the theory that INS was misappropriating material that AP put together (i.e. that INS was being unjustly enriched).

The basic idea is that compiling the news reports required a substantial input of time, money, and other resources, and so a second comer should not be allowed to compete with the compiler and profit from the compiler's work by simply copying the material.

(2) The Motorola Case: Significant clarification of (and limitation to) the INS case was made in a case referred to as the Motorola Case.[xiii] The defendant (Motorola, doing business as SportsTrax) was sending basketball scores and statistics to pagers with close to "real time" updates (for a fee, of course).[xiv] The Plaintiff (NBA) recognized that it had no property interest in the scores and statistics themselves but claimed that the material should be protected under the "hot news" theory developed in INS.[xv] The court disagreed and allowed Motorola to transmit the information.

The holding is significant because the court recognized that allowing INS type claims involves avoiding the statutory preemption requirements of the Act. Given that the claims are an exception to the Act (and so, presumably, narrowly construed), the Motorola court limited INS-like claims to cases where:

> (i) a plaintiff generates or gathers information at a cost; (ii) the information is time-sensitive; (iii) a defendant's use of the information constitutes free riding on the plaintiff's efforts; (iv) the defendant is in direct competition with a product or service offered by the plaintiff's efforts; and (v) the ability of other parties to free-ride on the efforts of the plaintiff or others would so reduce the incentive to produce the product or service that its existence or quality would be substantially threatened.[xvi]

Generally speaking, Motorola limited and defined the applicability of INS.

(3) The PGA Case: a more recent case that discussed (and distinguished) Motorola may be referred to as the PGA case.[xvii] Much of the PGA case revolves around anti-trust and monopoly issues, however, the case pulls out some important elements of Motorola. In the PGA case the Plaintiff (Morris Communications) wanted access to PGA real-time scores for free.[xviii] In its arguments regarding property rights, Morris analogized the case to Motorola for the premise that scores are not proprietary information and so are not protected (absent a hot news exception).[xix] The factual layout of the two cases, however, is not analogous and so the court refused to apply Motorola.

To make a long story short, the scores in the Motorola case were obtained by employees of Motorola who would watch an NBA game on television (or listen on the radio) and input the data as they viewed the event. In PGA golf events, however, the only way to get real time scores for all the players was to a special system that was developed and implemented by the PGA.[xx] This real time information was available at a command post on the tour grounds and certified members of the media were allowed access to the information as well as authority to publish the information.[xxi] The authority to publish, however, was limited by license and the scores could not be transmitted to websites or other broadcast media until 30 minutes after the shot or until the PGA made the information available to the public.[xxii] The second comer in PGA did not have access to all the real time scores without going through the PGA (i.e. the scores were not yet in the public domain).

In Motorola, however, the scores were in the public domain (through television or radio broadcast) and so were no longer proprietary information when accessed and re-published by the second comer. The bottom line comment on Motorola (by PGA), then, is that one need not give out compiled material for free, but, once it has been broadcast/published, the underlying information is no longer protected absent a hot news exception.

Policy Choices

Imitation of Copyright Law

One potential avenue for protecting databases involves imitating copyright law. This involves creation of federal legislation that provides databases with protection similar to that currently offered to qualified works under Copyright law. While this is undoubtedly attractive to parties who develop and maintain databases, it is questionable on several levels.

The Act is empowered by Constitutional language that reads, in relevant part, as follows: "Congress shall have Power…To Promote the Progress of Science and useful Arts, by securing for limited Times, to Authors and Inventors, the exclusive Right to their

respective Writings and Discoveries."[xxiii] It would be a stretch to find that compilation of a database qualifies one as an author or inventor and so extension or modification to the Act itself may run into constitutional problems. Therefore, legislation under this approach would presumably have to come from the Commerce Clause. Even so, the legislation would conflict with the Act in so far as the Act has explicitly addressed the issue of compilations and so the Act would have to be modified.

Extension of the Hot News Exception

A second potential avenue is to extend the protection available to certain "hot news" to certain databases. The databases entitled to protection under this theory would presumably be limited in scope similarly to the news that is entitled to protection. This may involve only protecting (1) time sensitive material, (2) that is developed at a cost to the first creator of the database; and (3) for a limited time only. This would seem to take into account and apply the justifications and theories developed in the hot news cases. Additionally, it may avoid the constitutional implications that creating a copyright-like regime would trigger.

Retaining the Status Quo

A third potential avenue is to simply leave things as they are (i.e. institute no new laws/extend no laws protecting databases. Whether this approach is prudent takes us back to the big policy question of "do we need additional protection for databases." Developers and managers of databases will generally say yes; schools and libraries will generally say no. The bottom line societal question becomes, if we leave things the way that they are, will companies lose any incentive to develop and maintain databases? If the incentive is lost, and the databases are not created, then society, as a whole, is losing out. However, as things currently stand, it is unclear whether there is truly a need for additional database protection. Potentially, the enactment of the EU Directive creates a need for database protection because absent database-specific legislation (that is equivalent to the EU Directive) U.S. databases will not be entitled to protection in countries that utilize the EU Directive.

Current Legislation Under Consideration

HR 3872 vs. HR 3261

The current database legislation pending in the "Consumer Access to Information Act of 2004" ("CAIA").[xxiv] The CAIA was introduced on March 2, 2004 and is an alternative to proposed legislation entitled "Database and Collections of Misappropriation Act of 2003" ("DCMA").[xxv] Under the proposed CAIA, the Federal Trade Commission would be responsible for oversight and enforcement of the act and private parties would not be authorized to sue under the title[xxvi].

The CAIA prohibits "misappropriation of a database" and goes on to define "misappropriation of a database" as follows;

> (1) a person (referred to in this section as the "first person" generates or collects the information in the database at some cost or expense;
> (2) the value of the information is highly time-sensitive;
> (3) another person's (referred to in this section as the "other person") use of the information constitutes free-riding on the first person's costly efforts to generate or collect it;
> (4) the other person's use of the information is in direct competition with a product or service offered by the first person; and
> (5) the ability of other parties to free-ride on the efforts of the first person would so reduce the incentive to produce the product or service that its existence or quality would be substantially threatened.[xxvii]

This definition is a more narrow definition then the DCMA has proposed. The DMCA would make it illegal to take a "quantitatively substantial" part of the information in a database and make it commercially available in the same market (without authorization) if:

> (1) the database was created and maintained through "substantial expenditure of financial resources";
> (2) the unauthorized use of it "occurs in a time sensitive manner and inflicts injury on the database"; and
> (3) the abilities of the unauthorized users to "free ride on the efforts" of the owner threaten the "incentive to produce the product" and consequently the existence of the database.[xxviii]

Under the DCMA anyone who either created or maintained a database would be protected from unauthorized use of the information in it. Opponents of the DCMA (largely schools and libraries) suggested that the DCMA utilizes ambiguous language (which would result in increased litigation); challenges (or eliminates) traditional "fair use" exceptions; does not allow for the transformative uses of information; does not include any "first sale" provision; does not include safeguards against monopolistic pricing; and raises fundamental questions about the relationship between the proposed legislation and copyright law. The end result, opponents fear, will be reduced competition in the database market and higher prices for libraries and schools.

Proposed Legislations Relation to Policy Approaches

Looking back to the potential avenues for database protection, the DMCA proposes protection that is similar to that of Copyright protection (perhaps even greater) while the CAIA proposes protection that appears to be codification of the hot news protection available under case law. Obviously, the CAIA is thought more favorably of by opponents of database legislation in general than is the DMCA. Although it is unlikely that either bill will pass, the CAIA marks a move towards more compromising

legislation; hopefully a sign that some type of resolution may be reached in the near future.

Conclusion

After reviewing Copyright Statues and case law as well as non-copyright (i.e. "hot news") case law, it is interesting to see the manner in which proposed legislation mirrors the legal history. Even much of the proposed language is taken directly from case law. There seems to be a huge push to get some type of legislation passed. That push presumes that there is a real and immediate need for such legislation. Given the explicit language found in the Copyright Act, as well as the Constitution, that question (whether there is a real need) is of paramount importance and the answer ought not be assumed.

References

[i] 17 USC § 103
[ii] 17 USC § 101
[iii] 17 USC § 102(b)
[iv] 17 USC § 103(b)
[v] Feist Publications, Inc. v. Rural Telephone Service Co., 499 U.S. 340, 111 S. Ct. 1282, 113 L. Ed. 2d 358 (1991).
[vi] Id.
[vii] Id.
[viii] 17 USC § 301(a) providing: "On and after January 1, 1978, all legal or equitable rights that are equivalent to any of the exclusive rights within the general scope of copyright as specified by section 106 in works of authorship that are fixed in a tangible medium of expression and come within the subject matter of copyright as specified by sections 102 and 103, whether created before or after that date and whether published or unpublished, are governed exclusively by this title. Thereafter, no person is entitled to any such right or equivalent right in any such work under the common law or statutes of any state."
[ix] See Harper & Row, Publishers, Inc. v. Nation Enter., 723 F.2d 195, 200 (1983), *rev'd on other grounds*, 471 U.S. 539, 105 S. Ct. 2218, 85 L. Ed. 2d. 588 (1985).
[x] International News Service v. Associated Press, 248 U.S. 215, 39 S. Ct. 68, 63 L. Ed. 211 (1918).
[xi] Id.
[xii] Id.
[xiii] National Basketball Association v. Motorola, Inc., 105 F.3d 841 (2d Cir. 1997)
[xiv] Id.
[xv] Id.
[xvi] Id.
[xvii] Morris Communications Corp. v. PGA Tour, Inc., 117 F.Supp.2d, 56 U.S.P.Q. 2d 1952, 28 Media L. Rep. 2544 (M.D. Fla. 2000); Morris Communications Corp. v. PGA Tour, Inc., 235 F. Supp. 2d. 1269, 31 Media L. Rep. 1642 (M.D. Fla. 2002); and Morris Communications Corp. v. PGA Tour, Inc., 364 F. 3d. 1288, 70 U.S.P.Q. 2d 1446, 32 Media L. Rep. 1513 (11th Cir. 2004) (all opinions contain discussion of Motorola).

[xviii] Id.
[xix] Id.
[xx] Id.
[xxi] Id.
[xxii] Id.
[xxiii] Article I, Section 8, Clause 8.
[xxiv] House Report 3872; Sec. 1.
[xxv] House Report 3261.
[xxvi] HR 3872; Sec. 4(a).
[xxvii] HR 3872; Sec. 2(b).
[xxviii] HR 3261; Sec. 3(a).

The Directive on the Legal Protection of Databases

Lead Author: Damon May

The European Union created The Directive on the Legal Protection of Databases (hereinafter the Database Directive; the Directive) in 1996. The Directive creates new rights for the owners of databases. It does this in two ways:

1. The Database Directive explicitly extends copyright protection to collections of data
2. It creates a new protection for databases as works in their own right. This *sui generis* protection aims "to prevent the unauthorized extraction and/or re-utilization of the contents of a database" [1]

These new protections offer a great deal of power to the owners of databases by placing severe restrictions on the use of those databases. Although the intent is to protect the investment of database owners, the protection has potentially damaging effects on scientists and other users of databases who have legitimate noncommercial need for the data.

Defining the Directive

Perceived Need for a Database Directive

Barring cryptographic or other precautions taken by their owners to protect them, databases "can be copied or accessed at a fraction of the cost needed to design them independently" [1]. The European Parliament and the Council of the European Union created the Directive out of a perceived need for increased protection for database owners.

This is true of many types of products in the information age. However, databases are mere collections of data and not original works in and of themselves, so they are not explicitly protected by any uniform copyright law throughout the EU. In the words of the Directive, "unharmonized intellectual property rights can have the effect of preventing the free movement of goods or services within the Community" [1].

The expressed fear of the Council, in the Directive, is that "[database] investment… will not take place within the Community unless a stable and uniform legal protection regime is introduced" [1]. The aim of the Directive is to provide a uniform environment in which all database makers can feel that their investment will be protected, thereby encouraging investment.

Rights Provided by the Directive

Copyright

The copyright protection afforded by the Directive is easily understood: “databases which, by reason of the selection or arrangement of their contents, constitute the author's own intellectual creation shall be protected as such by copyright” [1].

However, the authors of the Directive note that, even under copyright protection, the database maker is exposed to “the risk that the contents of his database may be copied and rearranged electronically, without his authorization, to produce a database of identical content” [1]. The copyright protection offered by the Directive can extend only to the *arrangement* of the data because the data can come from many sources and may already be copyrighted by other parties.

Sui Generis

The authors of the directive felt that simple copyright protection left database owners too exposed to infringement. The *sui generis* protection created by the Directive is an entirely different beast. It prevents unauthorized users from making use of the database in a way that interferes with the exploitation of the database by its owners.

Two things are specifically prohibited by the *sui generis* protection:

- **Extraction**: “the permanent or temporary transfer of all or a substantial part of the contents of a database” [1] (an explicit exception to this protection is made for extraction necessary for viewing of the database contents by an authorized user)
- **Re-utilization**: “making available to the public all or a substantial part” [1] of the database

The *sui generis* protection is intended to be quite broad: it relates “not only to the manufacture of a parasitical competing product but also to any user who… causes significant detriment… to the investment” [1]. The list of potential offenders is limitless – scientists, librarians, hobbyists, and anyone else who might inadvertently harm the investment of the database creator seem to have a good deal to fear from the *sui generis* protection.

Exceptions

Very few restrictions are placed on the protection described above. There is a brief nod to competitiveness concerns: “protection by the sui generis right must not be afforded in such a way as to facilitate abuses of a dominant position”. Most of the actual exceptions, however, are left to the member states to implement if they choose, and there are strict boundaries placed on those exceptions.

Members are allowed to create exceptions for “private purposes”, teaching or scientific research, public security, and administrative or judicial procedure. The Directive does

not mandate that the member states implement any of these exceptions. Furthermore, *all* of the exceptions potentially created by member states are trumped by "the rightholder's legitimate interests or conflicts with normal exploitation of the database" [1].

The specifics of the member country pieces of legislation implementing the Directive, then, are hugely important. These member states have varied a great deal in the exceptions they've included. Austria, Germany, and the UK have created all of the allowed exceptions for scientific and educational use. France and Italy, on the other hand, have not created any exceptions for science or education at all, and Ireland only allows such exceptions for non-electronic databases [16].

One highly important "exception" that is included in the Directive is that the *sui generis* protection only applies to nationals of a non-Member country if that country offers "comparable protection". This is clearly meant to goad other countries into adopting similar measures and creating a "uniform legal protection regime," not only within the EU, but across the globe and to give EU companies a competitive advantage against competitors in noncompliant countries.

Enforcement of the Directive

The Directive has been in effect for eight years now, and the last member countries enacted legislation enforcing the Directive four years ago. We haven't yet seen many interesting court cases yet dealing directly with the scientific use of data, but there have been some interesting cases with important implications for the future of the Directive. The battles are being fought in more mundane arenas, such as protection of telephone directories (Germany and France), self-help pamphlets (Belgium), and real estate advertisements (Germany) [4].

In many cases, the courts have held up the plaintiff's database right. In Germany, this has been the case in *Berlin Online* 1998, *Suddeutsche Zeitung* 1998, *Tele-Info-CD 1999, Kidnet/Babynet* 1999, and others. In The Netherlands, the plaintiff won out in *KPN vs. XSO* 2000. And in Spain, the right was upheld in *Editorial Adanzadi* 1999 [4]. These cases have certainly shown that the Directive has teeth.

Although none of these cases dealt directly with scientific use of data, there is every reason for scientists to believe that they could end up as an unsuccessful defendant in a similar case – especially in the countries where there are no exceptions in the member state legislation for scientific use. The deterrent effect of these decisions on scientific research cannot, of course, be measured, but there is every reason to believe that there has been such an effect.

One very important Database Directive case illustrates some new developments in enforcement of the Directive. The case was filed in the UK; decided for the plaintiff by the British High Court of Justice; appealed to the English Court of Appeal, which also found in favor of the plaintiff; and finally appealed to the European Court of Justice

(ECJ). On a parallel path, another plaintiff filed similar cases in three different countries, with very similar results and an eventual appeal to the ECJ for clarification.

In a decision affecting all of these cases, the ECJ recently reversed the earlier courts' decisions and found for the defendant, taking some wind out of the sails of the Directive.

British Horseracing Board v. William Hill Organization Ltd.

The British Horseracing Board (BHB) creates and maintains horseracing information – lists of races, horses, jockeys, etc. They license this information to bookies. William Hill Organization began making commercial use of BHB's information, obtained from BHB's website, in their own Internet service.

In July 2001, the BHB brought suit against the William Hill Organization, accusing the latter of infringing upon their database right. BHB contended that this daily use of information was an extraction or reutilization of a substantial part of the database and therefore was covered by the *sui generis* right established in the directive. Alternatively, they asserted, William Hill's periodic usage could amount to a systematic and repeated extraction of insubstantial parts of the database, also covered by the *sui generis* right.

The English High Court ruled in favor of BHB, finding that William Hill had undermined BHB's ability to exploit the database for commercial gain. William Hill appealed to the Court of Appeal, who agreed in principle with the High Court. However, they referred some key questions to the European Court of Justice (ECJ), notably "Is publicly-available information coverable" and "Do constant updates create a new right?" [9]

Fixtures Marketing Cases

Fixtures Marketing Limited manages the licensing of fixtures lists (scheduling information) for the English and Scottish Premier football leagues to companies outside of the UK. Three different bookies in Finland, Sweden, and Greece obtained this information and provided it to their customers.

The Fixtures Marketing cases resemble the BHB case in many respects. They differ somewhat in that, instead of obtaining the information from the plaintiff, the defendants obtained it indirectly through newspapers and other sources that had presumably licensed the information themselves.

ECJ Decision

Before the ECJ made their judgment in these cases, the EU's Advocate General, Christine Stix-Hackl, delivered her opinion on the issue. Agreeing with the British courts, in June 2004 she stated that William Hill and all three bookies had re-utilized information from the plaintiffs' databases in violation of the database owners' right.

Most observers with an interest in the case expected a similar ruling from the ECJ. And, in fact, on some of the questions put to them, the ECJ bolstered pieces of the Directive. For instance, they made it clear that infringement of the database right could occur whether the offending party accessed the data directly or, as in the Fixtures Marketing cases, through an intervening party.

Even more significantly, the ECJ found that the database owner's right is not diminished if the contents of the database are made public, even if it was made public with the owner's consent. So the fact that BHB made their information available on their website did not imply that William Hill had the right to make use of it however they saw fit.

However, despite these findings, the ECJ surprised a good number of people by finding in favor of the defendants on one key question. Their central argument is as follows: "The resources used to draw up a list of horses in a race and to carry out checks in that connection do not constitute investment in the obtaining and verification of the contents of the database in which that list appears." [10] The databases in question were not protected because the arrangement of the data did not represent significant investment on the part of their creators *separate* from the investment in creating the data.

It's a mixed decision in terms of support for database protection. On the one hand, the ECJ has made it clear that single-source databases such as these, in which the database owner is also the creator of the data, can be protected under the Directive, even if made public. On the other hand, however, they've made it clear that a database must demonstrate significant, separate investment in the arrangement of data in order to qualify for protection. This makes some progress toward settling a long-standing question about the Directive [12].

The ruling begs the question: under what circumstances could a single-source database be protected under the Directive? How could a database maker demonstrate significant investment in the arrangement of a database, separate from the creation of the data itself? This ruling seems to encourage abuse. Perhaps a future BHB or Fixtures Marketing could win a similar case simply by assembling its data in a subideal form and then rearranging the data into a different form and documenting that effort faithfully. No one would gain by that kind of wasted effort.

The ECJ also leaves a great deal of ambiguity around single-source databases by making their language specific to horseracing lists. Since database makers have great incentives to seek protection under the Directive, this invites similar cases with different types of single-source databases.

However, the ECJ decision does begin to dispel one fear related to database protection. As of October 2001, 50% of all the lawsuits had been brought by companies whose databases were incarnations of datasets that they themselves had created, such as telephone listings and broadcast schedules [17]. Many in the community feared that this use of the database right would give too much power to creators of these kinds of

artificial data. The ECJ's ruling, while not perfectly clear on this issue, goes a long way toward allaying these fears.

Responses to the Directive

Official Responses

The text of the Directive requires a review of applications of the Directive every three years after the implementation deadline of January 1, 1998. This review is designed to "verify especially whether the application of this right has led to abuse of a dominant position or other interference with free competition" [1] and potentially to lead to amendments to the Directive to prevent such interference.

The first such review would have been scheduled for 2001, but since only three member states met the 1998 implementation deadline the review has been delayed significantly. In the meantime there have been two major official analyses of the Directive. These analyses are significant since they will inform any potential changes to the Directive.

Nauta Dutilh

In 2002, the Commission requested an independent review of the Directive by Nauta Dutilh, an Amsterdam-based legal firm. Nauta Dutilh responded with some very real but narrowly focused criticisms of the Directive. Nauta Dutilh's criticisms center on the *sui generis* right – they claim that "the sui generis right has introduced a serious imbalance between the rights of database users and producers" [7].

Nauta Dutilh feels that libraries, in order to continue to compile and preserve data in the era of the Directive, are faced with an unreasonable administrative and legal burden. Since the actions of libraries are necessary for data preservation and are unlikely to have adverse effects on the database owners' investments, Nauta Dutilh suggests an exception for preservation of data by libraries, perhaps through mandatory licensing of data.

(Libraries have been particularly vocal opponents of *sui generis* protection, and they have received some sympathy in the EU. As early as 1997, an EBLIDA report singled out libraries as particularly vulnerable to the effects of *sui generis* protection because they acquire databases on a daily basis[1] [20].)

Though Nauta Dutilh has a lot to say about the need to protect libraries from the maleffects of *sui generis* protection, they fail to champion other affected groups, such as researchers.

Nauta Dutilh does point out a glaring ambiguity in the Directive with regard to database updation. If a database is being constantly updated, they ask, "Is the term of protection

[1] Contrast with BHB's concerns with their databases of daily races, above, and online matchmaking companies' presumed concerns regarding their databases of dating aces.

for the entire database renewed each time" [7], or does the extended term of protection only apply to updates? Nauta Dutilh suggests that the latter interpretation makes the most sense, and that this could be made possible with a datestamping requirement.

This, to me, seems an intractible problem. Individual database fields can be datestamped, true. But exposing that datestamp on every field in every manifestation of the database seems impossible. Nauta Dutilh's vague suggestion of "datestamping" as the cure for this particular quandary seems wholly inadequate. Overall, the Nauta Dutilh study seems patchy and its proposed solutions to the Directive's problems unsatisfying.

Commission Staff Working Paper

A Commission Staff Working Paper released this summer indicates that the first of the formal investigations of the Directive is now expected by summer 2005 [6]. The Working Paper offers a few hints of what may be to come. It indicates that the Database Directive may need some exceptions for the benefit of the disabled – the Directive has no exception for the nonvisual formatting of data.

The Working Paper's findings are based on the Nauta Dutilh study to a large extent. Like Nauta Dutilh, the Working Paper points out the need for an exception for preservation of data by libraries – this exception is provided in the EU's Information Society Directive but not explicitly in the Database Directive. However, the Working Paper states that "only minor adjustments seem to be necessary at the moment." [6]

Community Responses to the Directive

In the wake of the ECJ decision described above, many in the sports betting community have already responded voluminously and passionately. One article claims that the ruling could "jeopardise the future funding of British racing", and a lawyer specializing in gambling voiced the opinion that the decision would "limit the scope of racing to sell data… quite dramatically"[19]. Another journalist went so far as to declare that racing faces "an unprecedented crisis… [threatening] the future funding and modernisation of the sport" [18].

These fears come directly from the pocketbook of the racing industry. The sale of data to bookmakers is worth £100 million per year and is therefore a significant source of funding to the industry. The British racing industry had planned shortly to move from its current data distribution model to one based on the licensing of data, and the ECJ ruling will certainly require some reexamination of that move.

Many in the sports betting community see the Directive as a necessity. However, the data in question are undoubtedly protected by copyright, just as they always have been. The racing industry had hoped that the Directive would provide them a powerful, if blunt, new weapon for securing their funding. The ECJ ruling simply leaves them with the protection that they had before the Directive was enacted.

The National Research Council (NRC), an American group, has expressed its concerns about the Directive and its impact on science since 1996. They take special issue with the *sui generis* right: "the most borderline of all the objects of protection under intellectual property law… raw or factual data… paradoxically receives the strongest scope of protection available from any intellectual property regime except, perhaps, patent law" [13].

The NRC sees dire consequences for "the advancement of science, the growth of knowledge, and opportunities for innovation" [13] if the pressure created by the Directive drives the US or any international body to adopt a similar law.

The International Council for Science (ICSU) conducted a Workshop in 2000 to evaluate the effects of the Directive and to suggest ways to lessen its negative impact on science. The Workshop participants focused on the need for the Directive to define "fair use" more clearly and more favorably to scientists, and to allow re-utilization of data under fair use, not just extraction [14].

Most of the responses from the scientific community to date focus on potential maleffects of the Directive on scientific research in general. The common complaint is that, under the Directive, scientists may be burdened with higher administrative and economic costs for access to the data necessary to do their work [15]. The ICSU CODATA group uses the example of a widely used scientific publication on global warming (*Trends '93: A Compendium of Data on Global Change*) that relies on the availability of data from a large number of sources. Under *sui generis* protection as provided by the Directive, they assert, "it is almost certain that *Trends '93* would not exist " [5].

Evaluating the Directive

We've seen that the Database Directive has the potential to do a great deal of harm to the scientific community. Libraries, scientists, and others who rely on the ability to make use of data from a large number of sources have expressed concern that the *sui generis* protection established by the directive may cripple their ability to function. There are specific examples of important works that may be impossible to create with such protection in place and enforced. Clearly the Directive poses a threat to the scientific community.

It appears, however, that these concerns are beginning to register with the EU. The official investigation into the effects of the Database Directive commissioned by the EU has recommended that the mandatory licensing of data be used to ensure that data are not lost completely because of the actions of rights holders. The EU will take these recommendations into account when they conduct their formal review of the Directive with an eye toward amendment.

More concretely, the ECJ has recently surprised many different groups with a judgment that narrows the scope of the database efforts that are protected by the Directive. We will undoubtedly see more action in this area over the coming months, but the EU may be rethinking the amount of power that the Directive gives to database owners.

The concerns of groups like the NRC and the ICSU, while understandable, may not be as relevant if the ECJ and other high courts continue to rule in ways that defang the directive. Moreover, given the concerns expressed in Nauta Dutilh's report and the Commission Staff Working Paper, both of which will have an impact on the upcoming formal review of the Directive, some of these groups' gravest concerns may be addressed before too long.

The Database Directive most assuredly still has teeth, and it poses serious concerns for scientists. There is some hope, however, that some of the more dangerous pieces of the Directive will be eroded through court cases and amendment and that the EU will back away from its initial, extreme position.

References

1. Council Directive No. 96/9/EC, O.J.L 77/20 (1996); available at http://europa.eu.int/ISPO/infosoc/legreg/docs/969ec.html
2. Statutory Instrument 1997 No. 3032 (The Copyright and Rights in Databases Regulations 1997); available at http://www.hmso.gov.uk/si/si1997/1973032.htm
3. Federal Act Establishing the General Conditions for Information and Communication Services (1997); available at http://www.iid.de/rahmen/iukdgebt.html
4. P. B. Hugenholtz, "The new database right: Early case law from Europe" (2001); available at www.ivir.nl/publications/hugenholtz/fordham2001.html
5. "Threat to full and open access to data," Ferris Webster, CODATA Newsletter No. 65, August 1997; available at http://www.codata.org/data_access/threat.html
6. Commission Staff Working Paper on the review of the EC legal framework in the field of copyright and related rights (2004) ; available at http://europa.eu.int/comm/internal_market/copyright/docs/review/sec-2004-995_en.pdf
7. Database Directive 96/9/EC, European Commission Review (2002); available at http://www.eblida.org/topics/copyright/nautadutilh_aug02.doc
8. "UK Court of Appeal Rules on First Case Involving Online Database Rights". IP/TECH ADVISOR March 2002; available at http://www.goodwinprocter.com/publications/IPA_databaserightsUK_3_02.pdf
9. **"**DATABASE PROTECTION NARROWED: BRITISH HORSERACING BOARD v WILLIAM HILL**"**; The Simkins Partnership; available at http://www.simkins.co.uk/ebulletins/archive/TAFDatabaseProtection.aspx
10. JUDGMENT OF THE COURT (Grand Chamber) 9 November 2004 (1); available at http://www.curia.eu.int/jurisp/cgi-bin/gettext.pl?lang=en&num=79958890C19020203&doc=T&ouvert=T&seance=ARRET&where
11. "Press Release No. 89/04, 9 November 2004"; available at http://curia.eu.int/en/actu/communiques/cp04/aff/cp040089en.pdf
12. "Program Schedules, Event Data and Telephone Subscriber Listings under the Database Directive", P. Bernt Hugenholtz, 2003; available at http://www.ivir.nl/publications/hugenholtz/spinofffordham.html
13. "Private Rights and the Public Interest in Scientific and Technical Databases", National Research Council, 1999; available at http://cartome.org/nrc-db-intro.htm
14. "Proceedings of the Second ICSU-UNESCO International Conference on Electronic Publishing in Science", 2001; available at http://www.icsu.org/5_abouticsu/CDSI_web/EPS2/bo5afin.htm
15. "History of Database Protection: Legal Issues of Concern to the Scientific Community", Anne Linn, National Research Council, 2000
16. "Database Directive Evaluation Underway", P. Bernt Hugenholtz, International Council for Science CODATA Newsletter 86, 2003
17. "Europe's Database Experiment", Stephen M. Maurer, P. Bernt Hugenholtz & Harlan J. Onsrud, Europe's Database Experiment, Science, Vol. 294 (26 October 2001), p. 789-790
18. "Data verdict shock means BHB face crisis over funding", Richard Evans, 2004; available at http://www.news.telegraph.co.uk/sport/main.jhtml?xml=/sport/2004/11/10/shevan10.xml
19. "Surprise ECJ judgment in BHB/William Hill Case : no infringement of racing data", 9 November 2004; available at http://www.olswang.com/news.asp?page=newssing&sid=110&aid=753
20. "The EU Database Directive: Emanuella Giavarra reports", Information Europe" (issue 3, autumn 1997, p. 3-4)

Researchers' Reactions to HIPAA

Lead Author: Don Totten

Introduction

In this paper I explore the clinical researchers' views on statutory database protection in the form of the Health Insurance Portability and Accountability Acts' Privacy Rule, HIPAA. I then aim to reconcile the researchers' views on the subject with the costs and benefits of HIPAA implementation as he perceives them.

I have interviewed two physicians, an economist, and a research coordinator, all of whom are involved in conducting clinical trials at the NIAMS funded University of Washington Multidisciplinary Clinical Research Center. See **Appendix A** for notes from my interviews with these researchers.

HIPAA: Do the Costs Outweigh the Benefits for Clinical Research?

Each of the researchers I spoke with was able to articulate the aim of HIPAA[1] and professed a genuine commitment to the protection of patient privacy, yet the tone of each interview was negative. This disconnect between how researchers view the aim of HIPAA and the implementation of HIPAA is interesting. I contend that it has its roots in the belief that a statutory protection should only be enacted if projected benefits outweigh costs[8], and the disconnect arises from the perception that the costs of HIPAA greatly outweigh its' benefits.

To explore this contention I will first look at the researchers' perception of HIPAA in general.

How Do Researchers Feel About HIPAA?

I have asserted that the researchers I spoke with felt negatively about the implementation of HIPAA. Here are some examples taken from the interviews that illustrate my assertion:

> 1) The front line clinical researcher sees HIPAA compliance as yet another set of requirements which must be followed in order to get IRB[15] approval to conduct research[1].

2) Since IRBs had required that the researcher protect patient data prior to HIPAA, the researcher feels that there is no significant gain in patient privacy associated with the burden of complying with the new regulations[3].

3) Researchers think the rule is being enforced ambiguously. While he is being asked to go to extraordinary means to protect a patients' health information, the researcher sees a drug store pharmacist explaining aloud how a patient should take a particular drug well within ear shot of other customers[1].

4) Researchers feel that the rule is being interpreted and applied ambiguously by IRBs[3]. This is confirmed in much of the literature published to aid in the implementation of HIPAA. One security and privacy workgroup explains that: "It should be noted that some of the terms chosen in the final Privacy Rule have proven often confusing"[14].

Further exacerbating the problem is the fact that IRBs are charged with protecting the human subjects involved in trials at their institutions and protecting their institutions against legal action taken by or on the behalf of these same participants. There is no mandate for them to make the researchers life easier. When the rule has ambiguous interpretations it's in the IRBs' nature to choose the most restrictive interpretation[3].

5) The researcher feels HIPAA is being applied in areas it shouldn't be. One researcher suggested that, for example, HIPAA was intended to stop cases of systematic abuse and that it was not intended to keep physicians from discussing a patients status in the elevator or to prevent someone from accidentally seeing a patients chart as it lays open on a nurses desk[1].

It seems clear from my discussions with researchers that they are looking at the implementation of HIPAA negatively. I would say that this would not be the case if they saw significant gains in either patient privacy or his ability to conduct research, but unfortunately they don't. In the rest of the paper we'll look at some specific costs and benefits of HIPAA implementation as perceived by the researcher, weighing the costs against the benefits in an attempt to discover the source of the researchers' frustration.

What are the perceived costs of complying with HIPAA?

Researchers that I spoke with all seem to perceive the cost of implementing HIPAA as high. Making the problem worse is the fact that the costs are additive, or as stated in an NIH note: "It is important to note that the Privacy Rule does not replace or act in lieu of existing regulations"[10]. Here are some examples of costs as perceived by the researchers I interviewed.

- Small tasks can no longer be accomplished with informal agreements, now contracts and vendor relationships are required. These tasks had timelines that were measured in hours, now they are measured in weeks[3].

- Additional paper work which must address HIPAA requirements, as each institution has interpreted them, must be added to each study proposal[3].

- Study designers now choose to include fewer sites and less collaberation because of the time and effort associated with complying with the interpretations of HIPAA at each site[7].

- Patient participation in clinical trials may decrease because HIPAA requires additional authorization[3].

- Web servers that contain patient health information are being rounded up and placed in HIPAA compliant server farms[3].

- Institutions are sending all of their employees to HIPAA training[3].

- HIPAA compliance audits are being performed[3].

- Researchers are being asked to detail and defend database schema and redesign in cases where the schema is in conflict with HIPAA[3].

- Research centers are being moved into HIPAA compliant buildings[3].

Two specific examples were given that I feel highlight both the cost associated with HIPAA and the researchers' frustration with HIPAA implementation.

The first example deals with the removal of a traditional fixture on hospital wards the "clinic board". These boards hung prominently at the entrance of the ward so that physicians could see at a glance what room their patient was in. The "clinic boards" were removed because for HIPAA concerns[4].

The "clinic board" has been replaced by a desk, at the desk sits a nurse with whom the physician must now "check in" prior to entering the ward. The physician provides the nurse with his ID and the nurse provides the physician with a pointer to his patient. Unfortunately, hospital administrators haven't hired dedicated nurses for this role, instead the nurses on duty man the desk as available. It is often the case that physicians queue up to get access to their patients[4].

It's easy to see that this "check in" process would impose costs such as:

- slowing patient treatment
- adding a time burden to an already busy physicians and staff
- frustrating physicians possibly leading them to visit patients less frequently

Another example of HIPAA imposing a cost was given by a physician who prior to HIPAA routinely used email to send digital diagnostic images to colleagues seeking their help with image interpretation in difficult cases[4].

The images themselves contain patient identifiable data. For this reason sending them outside the institution without authorization and without a method of providing an audit trail of their use is not HIPAA compliant, and the practice has been stopped at an institutional level[4].

The research physician can still collaborate with his fellows and colleagues but the simple email process has been replaced with one that is much more time consuming and awkward. First the digital representation of the image must be edited to remove the patient identifiable information. Then the image is encrypted and placed on a secure web page. The physician can then email his colleague letting him know the user id and password of the page, and send a separate email with the encryption key. His colleague can then view the image and a recording of the fact that he has done so is automatically generated[4].

This is a cumbersome process as the physician must rely on a technician to do the lion's share of this work. This process adds real costs in the form of

- Delayed patient care
- Additional man hours
- Disincentive to collaborate
- Reduced quality of care[6]

In both these examples the costs could be troubling and the resulting incremental gain seems slight. Paradoxically these costs are exactly the kinds of costs the authors of HIPAA were trying to avoid as they say expressly that they wish to strike a balance between protecting patient privacy and allowing the transfer of information necessary to provide highest quality health care[2]. In the view of the physician interviewed, this balance was not achieved by removing the "clinic board" or restricting his email activities. I would conclude that the researcher feels as if HIPAA has imposed a fairly high cost as it has been implemented in his environment.

What are the perceived benefits?

We have just seen that the clinical researcher feels there some very real costs imposed on him by the implementation of HIPAA. I asked researchers if HIPAA had any direct benefits for them as researchers, all answered "no". For that reason I aim here to look for some subtle benefits which might not be apparent to the clinical researcher.

It has been suggested that one of the potential benefits of database protection might be the creation of new databases that do not exist today[8]. To explore this idea I asked researchers "what data isn't available currently that you would like access to?" One of the researchers I spoke with answered I'd like access to databases that enable me to

answer simple questions of the sort, “How many people had hip replacement surgery in the U.S. last year?”[1].

The researchers I spoke with were unaware of any new public initiatives that would create such databases[7]. One went further and stated that to his knowledge databases capable of answering these type of questions were not created primarily as a result of enabling technologies, standards or protections but rather had been created as a result of some monetary incentive[7].

There are some good examples that would tend to support his view. For instance submitting a claim to Medicare for reimbursement of the expenses associated with hip replacement surgery creates a “record” in the Medicare database. For this reason we can currently answer the question “How many Medicare recipients got hip replacement surgery last year.[7]?”

Another example would be Scandinavian countries. The researcher I interviewed suggested that many Northern European countries would know exactly how many of their people had hip replacement surgery last year because they have national health care systems. So, again the model of creating a national database as a byproduct of submitting reimbursement claims would apply here[7].

Another researcher mentioned that getting a list of people with asthma involved looking at clinic billing records[3] not patient medical records. These examples would seem to confirm the theory that treatment data is often only collected electronically as a byproduct of billing.

This isn’t entirely the case. The National Center for Health Statistics a publicly funded division of the Center for Disease Control compiles databases which represent the state of the nations health[11]. The physicians who contribute to these NCHS databases are selected by the NCHS but participate voluntarily. These databases are created with public dollars for the public good, not as a byproduct of billing. It’s interesting to note that the NCHS operates under privacy rules that are more stringent than those that HIPAA covered entities operate under[12], which would go to confirm the argument that statutory database protection enables the creation of databases.

Still I think the researchers point holds, monetary incentive has probably been the key factor in the success of such projects to date. The NCHS typically asks that physicians participate in the survey for a week[13]. To ask orthopedic surgeons to participate in a collaborative effort recording enough detailed information to enable us to answer our hip replacement question would require an ongoing commitment of time and labor which without incentive would seem unlikely.

My conclusion then is that it’s probably too early to tell whether or not database protection in the form of HIPAA has enabled the creation of new databases. It seems evident that there are other factors such as monetary incentive which play a large role in

the process of encouraging collaboration, and that we are unlikely to see new databases until all of these factors have been addressed.

Another area that we might find evidence of HIPAA providing benefit to the clinical researcher would be more or distinctly new forms of research collaboration. Collaboration in this environment would be aided by researchers being able to disseminate data more widely as a result of database protection[8].

To explore this theme I explained to researchers that my hypothesis was that HIPAA would lay the groundwork for open and free exchange of data between HIPAA compliant entities. Then I asked if each researcher could envision such a scenario and if they thought it could lead to something like open source clinical trials[9] in their environment.

The responses I got were favorable to theory of database protection enabling open source clinical trials[9]. However, the researchers I spoke with thought that database protection alone would not be able to address enough of the issues associated with conducting a multi site clinical trial to allow the idea to work in practice[4].

Again it may be too early to tell whether HIPAA has enabled new forms of research collaboration. Other factors may be confounding the creation of such collaborations and until they are addressed we may not see the benefit of HIPAA in this area.

It seems clear that researchers perceive little tangible benefit from the implementation of HIPAA. It seems equally clear that the more subtle network benefits of database protection have not yet made themselves apparent. On this front we do seem to have evidence that there are some confounding factors external to HIPAA which when addressed may allow these benefits to come to light.

Conclusion

Database protection as implemented by HIPAA is being seen by clinical researchers as adding to the price of doing business. Researchers understand the aim of HIPAA, but its associated implementation costs confound the already difficult task of conducting publicly funded clinical trials. It seems evident that, while HIPAA may provide some future benefit, the researcher sees little if any benefit currently. This has left the researcher with a profoundly negative opinion of HIPAA as, in his mind, the costs far exceed the benefits.

References

1. Interview with Dr. Richard Deyo MD, MPH Director NIAMS UW MCRC
2. Health and Human services web site, Summary of the HIPAA Privacy Rule - http://www.hhs.gov/ocr/privacysummary.pdf
3. Interview with Linda Levenson research coordinator NIAMS UW MCRC
4. Interview with Dr. Bevan Yeuh MD, MPH

5. Informal discussion with Patrick Heagerty PhD. Bioinformatics during this discussion Dr. Heagerty explained that during the grant renewal process his reputation was his currency.
6. The cost in terms of quality of care would be attributable to the fact that a physician may decide not to seek the advise of his colleagues because of his aversion to the process I've detailed.
7. Interview with Dr. William Hollingworth PhD
8. Science – Maurer and Scotchmer 284 (5417):1129
9. June 10 2004 Economist: An open source shot in the arm.
10. National Institutes of Health web site, note on IMPACT OF THE HIPAA PRIVACY RULE ON NIH PROCESSES INVOLVING THE REVIEW, FUNDING, AND PROGRESS MONITORING OF GRANTS, COOPERATIVE AGREEMENTS AND RESEARCH CONTRACTS - http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-025.html
11. National Center for Health Statistics web site, about NCHS - http://www.cdc.gov/nchs/about.htm
12. National Center for Health Statistics web site, How NCHS protects you privacy - http://www.cdc.gov/nchs/about/policy/confiden.htm
13. National Center for Health Statistics web site. NAMCS Description - http://www.cdc.gov/nchs/about/major/ahcd/namcsdes.htm
14. Work Group for Electronic Data Exchange, WEDI – Strategic National Implementation Process SNIP: http://www.wedi.org/cmsUploads/pdfUpload/WhitePaper/pub/P-Final-Notice_Auth.pdf Notice of Privacy Practices and Authorizations.
15. Institutions that conduct research which involves people have Human Subjects Review Boards or IRBs. They are responsible for the protection of the patient during the course of the research.

Factors in US Policy Decisions

Lead Authors: Elijah Esquibel and Charistel Ticong

The US government is currently faced with creating a balance between making certain that the rights of those producing databases are protected and at the same time not stifling the beneficial uses of scientific and technical databases. Many people in the database industry have expressed concern that they are out in the open, susceptible to "free riders" taking advantage of their hard work. The fact of the matter is, with relatively little or no technical knowledge and effort, data can be copied cheaply and effectively. This ease in copying and distribution of databases creates large disincentives for companies to invest in the creation of new databases and to horde those that they do create off line in the private domain.

The problem with creating a *sui generis* protection for databases is that there is a large amount of information in the public domain that can fall under the current term "a collection of data." For example, is my personal choice to arrange my music as I see fit a unique database. Is the organization within my iPod a protectible database? Likewise, is the database that Apple's iTunes site tracks and compares my purchases and the purchases of others, based on my collection, a protectible database? That information is important because, if someone's particular musical tastes are known, one can tailor advertisements to a specific group of people. The very definition of a database is not an original idea, it is rather the use of known facts, but those facts are "arranged for ease and speed of search and retrieval."

Typical Objectives of Organizations That Produce and Disseminate S&T Databases

	Federal Government	Not-for-Profit/Academic	For-Profit
Primary motivations	Serve national goals, including promoting societal well-being and supporting basic research and other public-good interests	Fulfill mission, including furthering research, education, creation of knowledge, and discovery; remain economically viable	Achieve corporate objectives, including profit making and growth, and ensure shareholder and customer satisfaction
Goals of S&T data collection and database	Support agency mission; undertake basic research as a	Advance knowledge by conducting new research and by validating and building on the research	Support development of new or improved products or services; develop databases for direct sale or

development	basis for economic growth and productivity and for public well-being	of others; educate future researchers; contribute to basis for producing social benefits; build reputation and status of researchers and their institutions	lease as products or as services in support of other products or services
Goals of S&T database distribution	Maximize the downstream benefits of basic research; promote availability and use of research results in both public and private sectors	Encourage open sharing of ideas; enable existing data to be reused for discovery of new knowledge; invite review and validation of research results; facilitate use of research results for product development by S&T community and commercial concerns; recover costs or generate revenue in support of mission	Disseminate data to protect competitive advantage when databases are used for development of other products or are themselves products or services; disseminate via sale or license to generate revenue, enhance customer base and market position, gain competitive advantage, achieve profits, or recover costs
Access to data	Open exchange of information encouraged by federal policy	Open, with data and ideas shared after results have been published	Internal and confidential, or available/marketed externally at a cost set by the organization
Interest in protecting the databases produced	Very low; any restrictions generally seen as a problem, with few benefits	Moderate; ranges from very low (for fully subsidized databases) to moderate (when cost recovery is necessary) to high (when data are a source of revenue required to support mission)	Very high; databases regarded as investments to be protected whether they are used in product development or are themselves products or services to be sold

[7]

Erosion of Copyright Law (a look at the aftermath of *Feist*)

1. Interpretation of Lanham Act
2. Click Licensing
3. Trespass to Chattels
4. Computer Fraud and Abuse Act

After the *Feist* case had subsided, many more cases have come to the surface hoping much to gain the much sought after protection of copyright law, but many have failed (well at least in the US). The potential consequences of enacted laws eroding the 1991 *Feist* case decision was evident in the *Dastar v. Twentieth Century Fox* case that was finally decided by the U.S. Supreme Court in (2003) that the Lanham (Trademark) Act could not be extended to include the attribution of facts. In the words of Justice Antonin Scalia "We do not think the Lanham Act requires this search for the source of the Nile and all its tributaries." Scalia goes on to include that "allowing a cause of action under §43(a) for that representation would create a species of mutant copyright law that limits the public's federal right to copy and to use" which in the case of database creation would make the whole premise that they are created for ease of interpretation and retrieval useless due to proposed attributions. These are the types of interpretations that uphold the Feist decision. The interpretations of the lower courts are what seek to reinstate the "sweat of the brow" defense that was rejected in *Feist*. The *Dastar v. Twentieth Century Fox* case was argued over many years before we were given a clear answer in 2003.

Shrink-wrap licenses introduced by product vendors of data in the form of CD's that will not allow access to the information purchased without agreeing to a click on license that appears during installation. The license in effect does evade *Feist* by holding the user of data to the agreement that they will not reuse facts contained within the product in any way shape or from. The case *ProCD* is where in which white pages are copied by Zeidenberg and in response the court has upheld the click-on agreement. This is a circumvention of *Feist* because white page listings are facts in the public domain that are now protected by this contractual agreement.

Trespass to chattels as in the *Ebay v. Bidders* Edge case show that a "mere interference with a possessory interest is sufficient to establish damage." The idea is that by searching Ebay's website with a web crawler Bidder's Edge was in fact inappropriately using Ebay's website to access information. It is due to this gross over use/ abuse of Ebay's conditional access to their database that resulted in a loss for Bidder's Edge. This idea of Cyber trespass can be used to protect facts not just the database itself. In effect protection of individual prices for items offered (facts) is what the court's ruling in the *Ebay* case has done. The court stood on the Intel v. Ham

The Computer Fraud and Abuse Act states that unauthorized attempts to upload information and/or change information on these web sites are strictly prohibited and are subject to prosecution under the Computer Fraud and Abuse Act of 1986 and Title 18 U.S.C. Sec.1001 and 1030. This is the wording of the U.S. Federal State department in that if a person in damage assessment as incurred al "loss" of $5,000 for even the extraction of one fact from his/her website or database the statue will go into effect. Two cases have used the CFAAU. In both *EF Cultural Travel v. Exploria* and *Register.com v. Verio* the defendants were found in violation of the CFAAU. Exploria was implicated for obtaining prices from EF so that they could undercut the competition and, Verio for accessing a personal database of customers of register.com in order to solicit to those patrons.

Congress would like to change all of this by allowing the first arranger to get control of it completely, including all the facts. In thinking about scientific data or publicly funded research the consequences are undoubtedly grim. Do we truly want research data frozen up like this? Facts that would normally enter the public domain will be owned privately, stifling further research that would otherwise be built on. Not to mention the scores of the NBA. There is no reason for this, and in fact AT&T, Yahoo, Google, Amazon, and many others contest to it, but the copyright alliance is strong, and their interests are awfully short-range.

Entities that support *sui generis*

1. Lexis-Nexis (Nigel Stapleton)
2. West/Thomson and Reed-Elsevier (Laura D'Andrea Tyson)
3. eBay (*eBay Inc. v. Bidder's Edge*)
4. Nasdaq (Robert Aber)

These are companies or businesses that rely on collecting databases and compiling them into an organized form. Without the protection of *sui generis,* their financial incentives are hindered by those that easily copy their databases, such as "free riders."[2]

According to a statement made by the chairman of Lexis Nexis, Nigel Stapleton, he stated that "competitors could potentially copy or extract significant portions of our databases and sell it in direct competition with us while avoiding the significant expense of creating it themselves."[9] Meaning that the fear is not much of the reproduction of the actual databases, but the fact that their very own databases could be used against them on the grounds of competition.

Laura D'Andrea Tyson who was the paid consultant for West/Thomson and Reed-Elsevier, and the Coalition Against Database Piracy (which is a coalition that was formed by West/Thomson and Reed-Elsevier) stated the same when it came to database protection. Tyson stated that "database providers spend hundreds of millions of dollars a year updating their existing databases and the newly-updated databases also need protection." Tyson stated that with the increasing technology there not only came improved access for the public to view databases, but easier access to duplicate and "steal" databases without "compensating" the creators of the databases. Tyson also suggested that the creators of the databases would not be able to make price differentiation when it came to the databases because if no protection was offered there would be no benefit because the database may be available for free. Tyson suggested to solve this was to add protection then charge "fees to be paid to the original authors and publishers." [10]

In April of 2000, eBay underwent a battle dealing with Bidder's Edge, a website that allowed on-line auction buyers to view a certain item that appeared in a number of auctions without having to search each auction individually, about "the method BE use[d] to search the eBay database." In the court case *eBay Inc. v. Bidder's Edge,* eBay alleged

that it would undergo incorrigible misfortune if Bidder's Edge continued to subsist and if relief were not granted. The damages it alleged were "(1) lost capacity of its computer systems resulting from to BE's use of automated agents; (2) damage to eBay's reputation and goodwill caused by BE's misleading postings; (3) dilution of the eBay mark; and (4) BE's unjust enrichment." Altogether, eBay alleged two categories of sufferings. "The first type of harm is harm that eBay alleges it will suffer as a result of BE's automated query programs burdening eBay's computer system ("system harm")." The second harm would be BE's misrepresentation of eBay's information therefore they would be suffering "reputational harm."[3]

In 1998, Robert Aber, who was the Senior Vice President and General Counsel of the Nasdaq Stock Market, Inc. and also the Chairman of the Board of Directors of the Information Industry Association ("IIA") made a statement declaring more protection. His concern for protection dealt with the competition with the international waters and the US as well. [1]

Aber feared that the U. S. database industry would be at a disadvantage with competition from other countries like Europe. He stated that the European Union's *Directive on the Legal Protection of Databases* ("EU Directive"), protected only the databases developed within its own country or other countries that had a similar laws enacted. Therefore, the EU Directive could be interpreted as a "license to steal" databases from countries overseas that did not recognize database protection. Since the U. S. produces the majority of the databases, its control and incentive might cease granted situations like this would occur. [1]

Aber also feared that there would be destruction within the US. Before the revelation of the *Feist* case, developers of databases believed that their databases were covered by copyright protection and sanctioned under the "sweat of the brow" doctrine. [1] Many believed that this doctrine would "prevent the copier from competing unfairly with the compiler by appropriating the fruits of the compiler's efforts or creativity. In this sense, courts treated copyright protection for compilations much like a branch of unfair competition law."[13] There was also fear that the compilation that has been created would not "express originality" and would thus retain limited protection from competing entities who want to make a quick buck off of someone else's hard work. Even if there was not a financial incentive, Aber worried that "cyberpranksters" would duplicate and display for public viewing on the internet without consideration for the producer of the database.

Entities that oppose *sui generis*

1. Researchers
2. Universities

There are various opponents/ skeptics to creating sui generis rights for databases. The issue of continuing research is what drives progress in the physical and social sciences.

Thus research expands our knowledge of the inter workings of complex systems within the world around us. Society uses this research in a wide varety of applications to increase the public welfare and create more wealth. The hording of facts through constraining sui generis protections would greatly disadvantage the practice of information exchange that occurs that this level. As Stephen M. Mauer proposes “Open Science” could be the key to saving millions of lives by the simple collaboration of researchers and doctors of third world diseases in that diagnosis, pharmaceuticals, and treatments in medicine techniques could be shared and learned from through “Computational drug discovery which is similar to de-bugging Software.” [5] In this respect agricultural researchers can also collaborate on better high-output food production and faster and cheaper energy production and manufacturing. NOAA has grown exponentially its databases from 250 mega bites in its first year of text and tables, to the projected 230 petaterabites in 2010. “Comprehensive database protection would turn the situation on its head by making virtually all facts protectible as ‘organized collections of information.’ ” [8] An imposition of a such sui generis protection would implore NOAA to take measures to restrict access to its vast collections of satellite images and raw data. This position is held by large companies like “Amazon.com, AT&T, Comcast, Google, Yahoo, and The US Chamber of Commerce [6]

A *sui generis* database protection on databases could lead to some unbelievable things such as a student copyrighting their own homework and a teacher would not be able to alter it unless the student gave permission to do so, sports quotes would not be able to be “quoted” unless granted, stock exchanges would be able to copyright “quote tables” and charge for viewing them, and last and oddly least just about anyone could copyright search engines and charge for the usage. [14]

Databases in the wrong hands (security and protection)

An example of internet community’s response to the unlawful theft of a database can be seen in the story of Half Life 2’s source code being stolen. “It was extraordinary to watch how quickly and how cleverly gamers were able to unravel what are traditionally unsolvable problems for law enforcement related to this kind of cyber-crime." [11] In this instance the stolen data was so damaging that it threatened to put the company Valve with a significant loss. The fact of the mater is there is a certain point when the community will recognize hard work and bring those responsible to justice. That standard of harm is in most cases too high for those who are for database protections.

University of California, Berkeley was ‘hacked’ and about 1.4 million people’s critical information on the university’s record keeping system was viewed. There is no way to tell if the data was compromised or accessed, but University officials say that “The information could potentially be used for identity theft or credit card fraud” [4] In this case anyone found to have accessed the database would be charged with unlawful use and access to the University’s database system and invasion of privacy by those whose information was captured by the hackers.

1. Aber, Robert E. "Hearing on H.R. 2652, the *Collections of Information Antipiracy Ac.t*" February 12, 1998 <http://www.hyperlaw.com/aber.htm>
2. Arnold, Giannia J. and Uri Bilek. "Databases Seek Legislative Shield." Oct. 11, 2004 <http://www.ebglaw.com/article_1049.pdf>

3. "EBAY, INC., *Plaintiff,* vs. BIDDER'S EDGE, INC., *Defendant."* April 14, 2000 <http://pub.bna.com/lw/21200.htm>
4. King, Rachel. "Campus Computer Hacking Spurs Federal Investigation." *The Daily Cal* October 20, 2004 <http://www.dailycal.org/article.php?id=16605>
5. Maurer, Stephen M. "Finding Cures For Tropical Diseases: Is Open Source an Answer?" <http://salilab.org/pdf/136_MaurerBIOESSAYS2004.pdf>
6. McCullagh, Declan. "Court Doesn't Extend Database Protection." February 26, 2004 <http://news.com.com/2100-1024_3-5165624.html?part=rss&tag=feed&subj=news [h] http://www.state.gov/documents/privacy.cfm>
7. National Academy of Sciences. "A Question of Science: Private Rights and the Public Interest in Scientific and Technical Databases." 1999 <http://books.nap.edu/html/question_balance/ch2_t1.html>
8. National Academy of Sciences. "A Question of Science: Private Rights and the Public Interest in Scientific and Technical Databases." 1999 <http://books.nap.edu/html/question_balance/ch3.html>
9. Stapleton, Nigel. "Reed Elsevier on Database Protection." June 1997 <http://www.cni.org/Hforums/cni-copyright/1997-02/1303.html>
10. Tyson, Laura D'Andrea. "Database Protection Proposals Page." September 29, 1997 <http://www.hyperlaw.com/dbpage.htm>
11. *Thorsen, Tor. "Valve Announces Half-Life 2 Code Theft Arrests." June 10, 2004* <http://www.gamespot.com/news/2004/06/10/news_6100381.html>
12. University of Maine. "Database Protection Laws Could Threaten Economic Development and Scientific Research." October 2001
13. U. S. Copyright Office. "Report on Legal Protection for Databases." August 1997 <http://www.copyright.gov/reports/dbase.html>
14. Wylie, Margie. "Crisis Over Copyrights." December 13, 1996. <http://news.com.com/Crisis+over+copyrights+-+page+2/2009-1023_3-254879-2.html?tag=st.num>

Conclusion

We've examined a number of forms that database protection has taken and may take in the future. Each of these implementations has its good and bad points.

In the United States, the various avenues for database protection represent varying degrees of protection. Copyright-like protection offers the most drastic protection while extension of "hot news" case law offers less protection than does imitation of copyright but provides more specific coverage than the existing law. The two most current proposed bills in Congress each represent a specific approach considered above. HR 3872 embodies the extension of the hot news case law while HR 3261 represents the more drastic, copyright-like protection. Not surprisingly, it seems that HR 3872 (the less drastic approach) is gathering more support.

The European Union's Database Directive has come under fire from many groups. Libraries and scientists fear that the Directive prevents them from doing their work. Under the Directive as it stands today, these concerns are well founded. However, recent case law has given these groups some hope for the future. Also, some flaws in the Directive will likely be addressed when the EU meets formally in 2006 to discuss amending the Directive. Though the Directive currently presents some very real concerns, some of its most drastic provisions may not be around forever.

Clinical researchers see database protection as implemented by HIPAA as an additional cost of doing business. Ethical researchers understand the aim of HIPAA, but its implementation confounds the already difficult task of conducting publicly funded clinical trials. They feel that protections in place prior to HIPAA adequately protected patient privacy and that HIPAA's additional regulations have little real value. While HIPAA may provide some future benefit, the researcher sees little if any benefit currently. In the mind of the researcher, the costs of HIPAA far exceed the benefits.

There are many entities such as Lexis Nexis and the Nasdaq who are in favor of *sui generis* protection. There are also those such as Google and Amazon.com who are against it. Both sides have their pros and their cons, but where to draw to that line of protection and availability is still unclear. On the one hand, the benefit of easy accessibility would allow for advancement in "Open science"; on the other hand, more protection would decrease chances of databases from being misused.

The opposition to the most extreme kinds of database protection is making itself heard in Europe, and support for such strong protection may be weakening somewhat there. At the same time, proposed legislation may bring the United States to a situation much like that of Europe, whether such new legislation is really necessary or not. In this very new area of government intervention, each country is still in the process of finding the right amount of power to place in the hands of the owners of databases.