

September 1997

A Thoughtful and Practical Analysis of Database Protection under Copyright Law, and a Critique of Sui Generis Protection

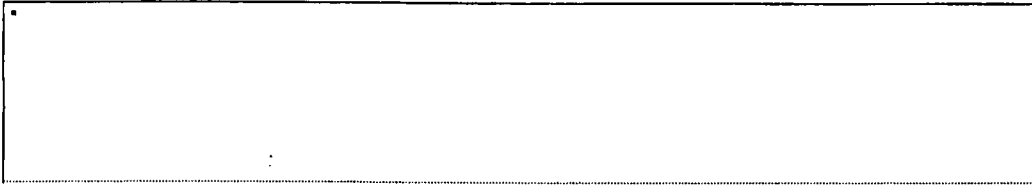
Wesley L. Austin

Follow this and additional works at: <https://scholarship.law.ufl.edu/jtlp>

Recommended Citation

Austin, Wesley L. (1997) "A Thoughtful and Practical Analysis of Database Protection under Copyright Law, and a Critique of Sui Generis Protection," *Journal of Technology Law & Policy*. Vol. 3: Iss. 1, Article 4. Available at: <https://scholarship.law.ufl.edu/jtlp/vol3/iss1/4>

This Article is brought to you for free and open access by UF Law Scholarship Repository. It has been accepted for inclusion in *Journal of Technology Law & Policy* by an authorized editor of UF Law Scholarship Repository. For more information, please contact kaleita@law.ufl.edu.



[Return to Table of Contents](#) [Comment on this Article](#)

A Thoughtful and Practical Analysis of Database Protection under Copyright Law, and a Critique of *Sui Generis* Protection

by Wesley L. Austin

Cite as: Wesley L. Austin, *A Thoughtful and Practical Analysis of Database Protection under Copyright Law, and a Critique of Sui Generis Protection*, 3 J. TECH. L. & POL'Y 1 <<http://journal.law.ufl.edu/~techlaw/3-1/austin.html>> (1998).

TABLE OF CONTENTS

I. Introduction

II. Standards Applied to a Database Containing a Compilation of Facts

- A. *Feist*: The Standard for Obtaining Copyright Protection
- B. Selected Applications of *Feist*

III. When the Rubber Hits the Road: A Practical Analysis of a Real Database

- A. A Conceptual Model of a Typical Database
 - 1. The Different Elements: the Computer Program and the Database
 - 2. Databases 101
- B. The *Feist* Standard Applied to a Database
 - 1. Original Selection
 - 2. Original Arrangement
 - a. The Physical Arrangement of the Database
 - b. The Logical Arrangement of the Database
- C. The Vulnerability of Databases
 - 1. Selection
 - 2. Arrangement
 - 3. Is This a Bad Thing?
 - a. The Free Rider Problem

IV. Attempts to Fix the Problem

- A. The European Directive
- B. The Database Treaty Considered at WIPO
- C. H.R. 3531
 - 1. Statutory Authority
 - 2. Critique of H.R. 3531
- D. Some Suggestions for the Next Database Protection Bill

V. Conclusion

I. Introduction

{1} A factual compilation can be protected by copyright law if the selection, coordination, or arrangement of data constitutes an original work of authorship. The facts themselves are not copyrightable. If the factual compilation qualified for copyright protection, the protection would extend only to the selection, coordination, or arrangement that made the compilation original. Protection would not extend to the facts contained in the factual compilation. As a result, the facts in a factual compilation may be freely copied. With the computer revolution, many factual compilations are taking the form of computerized databases. With the ease of copying electronic information, "free riders" may take a first database creator's database, copy the uncopyrightable elements, and make a second competing database without incurring the cost of producing it.

{2} This article analyzes the current copyright protection available for factual compilations in the database context. The focus is on databases containing uncopyrightable factual material. In other words, an assumption of this paper is that the data in a database is not copyrightable. Those databases that contain copyrightable materials, e.g., articles from newspapers, magazines, etc., will not need to reach the issues examined in a factual compilation case because the elements of the database will be copyrighted. [1]

{3} This article discusses the basic structure and operation of a database. Characteristics specific to databases are considered when analyzing the selection and arrangement of data in a database, and the selection and arrangement standard is applied to a hypothetical database.

{4} Finally, current *sui generis* legislation being considered in the U.S. and international initiatives that focus on database protection are reviewed. Several shortcomings of H.R. 3531 (the bill introduced last year in congress but not passed) are pointed out, and the substance of several new provisions is suggested.

II. Standards Applied to a Database Containing a Compilation of Facts

A. Feist: The Standard for Obtaining Copyright Protection

{5} The Copyright Act extends protection to original works of authorship. [2] A database of factual material falls under the statutory category of compilation. A compilation is a "work formed by the collection and assembling 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." [3] A copyright in a compilation "extends only to the material contributed by the author of such work, as

distinguished from the preexisting material employed in the work" [4]

{6} In the landmark decision of *Feist Publications, Inc. v. Rural Telephone Service Co.*, the Supreme Court indicated that copyright protection for a factual compilation is 'thin' and that the protection extends only to original selections, coordinations, or arrangements. [5] At issue in *Feist* was a white- and yellow-pages telephone directory. *Feist* used Rural's white pages listings without its consent. [6] Rural brought a copyright infringement suit against *Feist*. [7] The Supreme Court held that the listings were not protected by copyright [8]. In its opinion, the Court rejected the "sweat-of-the-brow" theory as a basis for copyright protection. [9] Moreover, the Court underscored that "no one may claim originality as to facts," [10] and "raw facts may be copied at will." [11]

{7} The court concluded that the telephone book's arrangement was mechanical and typical in deciding that its selection and arrangement were entirely devoid of creativity and undeserving of copyright protection. [12] In *Feist*, the item at issue was a telephone directory, in a hardcopy.

{8} Many subsequent cases involving factual compilations have involved hardcopy of some sort. These cases are especially useful when looking at similar factual compilations. However, an analysis of a database that is more than just an electronic copy of a book is much more challenging, as will be shown in section III.

B. Selected Applications of *Feist*

{9} In *Bellsouth Advertising & Publishing Corp. v. Donnelley Info. Publishing, Inc.*, the Court of Appeals for the Eleventh Circuit found a yellow pages directory not copyrightable. [13] Bellsouth Advertising & Publishing Corporation ("BAPCO") published a yellow pages directory. [14] The BAPCO directory was arranged alphabetically by business classification. [15] Donnelley prepared a competing directory by giving copies of BAPCO's directory to a data entry company that created a database containing the name, address, and telephone number of the entity. [16] "From this database Donnelley printed sales lead sheets and subsequently prepared its own competitive directory." [17] The court held that Bellsouth's selection, coordination and arrangement did not display the originality required to merit copyright protection. [18] The arrangement of the yellow pages was like the arrangement of the white pages in *Feist* and entirely typical.

{10} In *Key Publications, Inc. v. Chinatown Today Publishing Enterprises*, however, a panel of the Court of Appeals for the Second Circuit found the yellow pages of a Chinese-American directory copyrightable. [19] The directory included an arrangement of over 260 categories and 9000 businesses listed within those categories. [20] The court determined that the selection was original because the compiler exercised subjective judgment in selecting which businesses would be of greatest interest to the Chinese-American community. [21] The court also found the arrangement original because it was in "no sense 'mechanical' [22] and involved creativity on the part of [the compiler] in deciding which categories to include and under what name." [23] The court held that there was no copyright infringement of a competitor's directory that had 28 categories and 2000 listings. [24] The court stated that "[n]o substantial categories and their listings have been taken wholesale from the Key Directory," and that "[t]he organizing principles of the two directories are thus not substantially similar." [25]

{11} In *Skinder-Strauss Associates v. Massachusetts Continuing Legal Education*, a portion of a compilation containing a listing of attorneys and judges was found not copyrightable under *Feist*. [26] Skinder-Strauss published the Red Book, a compilation that included a directory of attorneys and

judges. [27] The Red Book listed the judges alphabetically by jurisdiction and included the name, address and telephone number of each judge. [28] The attorney listings both directories were arranged alphabetically and geographically. [29] Each attorney listing provided a name, telephone number, firm name, year of bar admission, address, and fax number. [30] Skinder-Strauss selected to list those attorneys who actively practiced in one of the six New England states. [31] In 1993, Massachusetts Continuing Legal Education published the 1994 Blue Book including much of the same information. [32] The district court held that Skinder-Strauss did not exercise a minimal degree of creativity in a Feistian sense in its listing of attorneys and judges, and accordingly, this portion of the Skinder-Strauss Red Book was not copyrightable. [33] The court also held that the merger doctrine applied because there are so few ways of compiling listings of attorneys, and therefore, a finding of substantial similarity under these circumstances was precluded. [34]

{12} In *Warren Publishing, Inc. v. Microdos Data Corp.*, the Court of Appeals for the Eleventh Circuit held that a publisher's selection of communities used in arranging its listing of cable television systems was copyrightable as a compilation. [35] Warren annually publishes a cable television factbook. This directory arranges entries alphabetically by state, and within each state, alphabetically by the name of the "principal" community served by the particular cable system. Information on each cable system is broken down into a uniform set of data fields. Warren contended that the elements copyrighted and infringed were (1) the communities covered, (2) the selection, sequencing and arrangement of the data fields, and (3) the content of the data fields. The court held that the selection of communities was copyrightable and that a software package infringed that copyright. Judge Kravitch pointed out the similarity of *Bellsouth*, and accordingly, dissented. The majority's opinion was vacated and a rehearing *en banc* was granted. [36] Perhaps others will see similarity between *Warren* and *Bellsouth*.

{13} These representative cases of *Feist's* progeny all involved plaintiff's factual compilations in book form. As a result, the order and arrangement at issue was a fixed aspect of the work and easily tangible. One could analyze the order and arrangement of the work by thumbing through the pages. When pages aren't there, the order and arrangement may get a little fuzzy.

III. When the Rubber Hits the Road: A Practical Analysis of a Real Database

A. A Conceptual Model of a Typical Database

{14} Before *Feist* can really be applied to a database, a basic understanding of how databases are usually structured is needed. After all, to determine if the selection, coordination, or arrangement is sufficiently original to warrant copyright protection, the selection, coordination, and arrangement of a database should be understood. Probably everyone understands what a phone book is, and most copyright practitioners understand how to apply *Feist* to such a work. [37] A trap that some may fall into is to just think of a database as an electronically stored book. Although one may look at a database this way, such a conceptual view distorts what a database really is, and effectively will distort any copyright analysis of the database. Hopefully after this brief explanation of a typical database, it will be clear that the similarities between databases and hardcopy factual compilations (i.e., books) are few and far between.

1. The Different Elements: the Computer Program and the Database

{15} As shown in Figure 1, a basic illustration of a user accessing a database may include three

different items: the user, the program used to access and search the database, and the database. The program usually provides a user interface and searching capabilities to help the user access information contained in the database. The user directly interacts with the program, and the program directly reads data from the database. Although Figure 1 only shows the database being read from, some programs allow a user to write to the database as well. However, because most databases do not allow users to write to them, Figure 1 shows the database as read only. In a typical situation where a user wants to access a database contained on a CD-ROM, the user first places the CD-ROM in the CD-ROM drive and then executes [38] the program that facilitates access to this database. This software is not the database, [39] but allows a user to search, navigate and view the database. Computer programs are treated as literary works under copyright law. [40] The important point to be made, in the context of database protection, about the computer program accessing the database is that the program has its own standard of copyrightability and should not be confused with the database itself. Therefore, a court should take care in analyzing the copyrightability of a database that it does not use part of the computer program in determining the copyrightability of the database. [41]

{16} Programs have user interfaces. User interfaces include such things as menus, toolbars, status bars, and windows. Anything created by a computer program that a user sees or interacts with in interfacing with the program is included as part of the user interface. User interfaces have a standard for copyrightability that is different from the standard applied to computer programs. [42] Again, the important point to be made is that the user interface, like the computer program, has its own standard of copyrightability and should not be confused with the database and its standard of copyrightability. [43] Courts must distinguish between the selection or arrangement of a database and the selection and arrangement portrayed by the user interface.

{17} Considering the foregoing, courts would do well to take an initial step in analyzing a database of uncopyrightable facts: turn off the computer, and leave it off. No matter how tempting it may be to want to look at the computer screen to see some selection and arrangement, as pointed out, anything on the screen has been manipulated and generated by the computer program and its user interface. For the copyrightability of a factual database, only analyze the database.

2. Databases 101

{18} The following explanation of a database is meant to be an aid to understanding how databases are structured. It is not the only way that databases are viewed, but it is a fairly common conceptual view used by software developers and database creators. The examples and explanations set forth are greatly simplified to focus on the legal aspects of database protection and avoid a technical course on database design.

{19} Databases usually include one or more tables of data, [44] as shown in Figure 2. Generally, each table is dedicated to holding many different items of

a certain type of information. For example, one table may be dedicated to holding individual's names, addresses, and telephone numbers, while another table may contain company names, addresses, telephone numbers, and fax numbers. If one database distributed on CD-ROM had three tables associated with it, there may be three data files on the CD-ROM, each corresponding to a table (e.g., table1.dat, table2.dat, and table3.dat).

{20} Each table usually contains a number of records. Each record usually has the same structure, but the information it contains is most likely different. As an example, if a database contained a directory of engineers, it would probably have a table containing the engineers directory. In this table would be a series of engineer records.

{21} Each record includes 1 to M fields, as best shown in Figure 3. Additionally, there may be 1 to K records in each table. An example of a table and its corresponding records will most readily convey the structure and format of a table and of the database. Let us examine a hypothetical database: the SmallTown Engineers Database.

SmallTown Engineers Database

Table 1: EEs

Joe
 USU
 BS
 joe@abc.com
 trekkie
 Betty
 UCLA
 MS
 betty@def.com
 trekkie
 Linda
 ASU
 BS
 linda@ghi.com
 trekkie

Table 2: SEs

Shawn
 UofU
 BS
 shawn@jkl.com
 trekkie
 Mary
 UofA
 Phd
 mary@mno.com
 trekkie
 Brent
 MIT
 BS
 brent@pqr.com
 trekkie

Table 3: MEs

Rita
 USU
 Phd
 rita@stu.com
 non-trekkie
 Jana
 BYU
 BS
 jana@vwx.com
 non-trekkie
 Jack
 USU
 Phd
 jack@yzz.com
 trekkie

Table 4: CEs

Heidi
 BYU
 MS
 heidi@abc.com
 non-trekkie
 Gary
 WSU
 BS
 gary@def.com
 trekkie
 Allen
 USU
 BS
 allen@ghi.com
 non-trekkie

Table 1. Selection and Arrangement of Hypothetical Database

{22} The SmallTown Engineers Database puts at the consumers fingertips a comprehensive listing of engineers in the city of SmallTown. For each engineer the database contains a name, the university attended, the degree obtained, and an email address. With hopes of obtaining copyright protection for the database, the database creator also included another piece of information: whether the engineer is a Star Trek fan.

{23} Table 1 illustrates the selection and arrangement of the SmallTown Engineers Database. As shown, the four columns in Table 1 each represent the data contained in one of the data files, and the arrangement of the data in each table. Each table consists of a plurality of records. Each record has the same format consisting of five fields: a name field, a university field, a degree field, an email address field, and a Star Trek field.

{24} Table 1 may depict the physical arrangement of the database: how the data is really arranged on the storage device. [45] Alternatively, it may depict a higher-level arrangement that will be referred to as the logical arrangement. [46] If an arrangement being described actually refers to how the data is physically arranged on some storage device, it is the physical arrangement. If not, it is a logical arrangement. This means that a user can't really tell if an arrangement is physical or logical unless the actual physical arrangement of the data on the storage device is known.

{25} The database also comes with a computer program for searching and displaying the SmallTown Engineers Database. With the computer program, a user can view the database arranged in many different ways. Through the computer program the user can view and/or create many different logical arrangements. For example, if the SmallTown Engineers Database was not physically arranged as shown in Table 1, but the access software allowed the user to view the database as shown in Table 1, the user is viewing a logical arrangement through the access software. The program may also allow a user to create a logical arrangement that was not there before. As an example, the program may allow the user to specify that the database of all the engineers appear in alphabetical order. The computer program can create this logical arrangement by reading in the data, sorting it, and displaying it to the user in alphabetical order.

{26} The way that the computer program actually displays the information to the user (how it appears on the screen) depends on the user interface. Some programs allow users to customize what is shown on the screen and how it is displayed. The computer program used to search the SmallTown Engineers Database could allow a user to select preferences such that only the name and university were displayed on the screen, and that these fields were displayed side by side.

{27} In operation, a user may access the database in the following way. Assume that a user wishes to search the database for any engineers with bachelor's degrees and only wants to view the results' email address fields. To accomplish this, the user first enters into preferences, a part of the computer program, that he only wishes to view the email address of records displayed. Then the user enters a search for all engineers with a BS degree and clicks "OK". The computer program then takes this input and performs a search on the database by searching through each table for BS in the degree obtained field and remembering where all these records were. When the search is completed, the computer program displays a portion [47] of the results through the user interface, which is what the user actually sees on the screen. The user then sees an arrangement of engineers' email addresses of all the engineers in the database with BS degrees. It is important to note that we have passed through at least three different works, each of which is protected under different copyright standards, in using the database: the

computer program, the user interface, and the database itself. With this crash course of databases 101, we are ready to move on to *Feist*.

B. The *Feist* Standard Applied to a Database

{28} As stated, *Feist* set out the rule that for a factual compilation to deserve copyright protection it must have an original selection, coordination, or arrangement that satisfies the minimum standards for copyright protection. [48] In other words, the selection, coordination, or arrangement must show a modicum of creativity to be copyrightable. [49] The selection part of the test will be analyzed apart from the coordination or arrangement part. The coordination or arrangement part of the test will be discussed together because it is unclear what the difference between coordination and arrangement is. [50]

1. Original Selection

{29} As applied to a computer database, the selection of facts is best applied to the selection of fields contained in a record, and the selection of records. This analysis can be applied much like the analysis used in case law involving hardcopy factual compilations because analyzing "what" is in a factual compilation really is not affected by whether the factual compilation is in book form, or is in electronic form. Thus, if the selection of fields or records is sufficiently original such that a court would conclude there was a modicum of creativity, the selection would be a copyrightable expression. The larger the quantity of data from which the selection was made, the more likely a selection can be protectible. [51]

{30} To solidify what selections we are talking about, an application of an original selection test to the SmallTown Engineers Database may be useful. For our selection of fields in each record (i.e., what bits of information we selected about each engineer), we have name, university, degree, email address, and whether they like Star Trek. From the directory cases we have seen, it may be assumed that none of that selection is original except the Star Trek field. A court would probably find that a minimal degree of creativity was shown in selecting that piece of information for compilation. For the selection of records (i.e., selection of what engineers to include in the database), let us assume that the database creator included every engineer that he could get information on. With this assumption, it seems clear that the selection of engineers included was not original because the database creator attempted to include as much of the universe of engineers as possible.

2. Original Arrangement

{31} It is this part of the *Feist* analysis that may get confusing if care is not taken to ensure that what one is looking at is, indeed, what one should be looking at. "Coordination and arrangement globally refer to the ordering or grouping of data." [52] However, despite this simple statement, the meaning of what an arrangement in a database means may be ambiguous. [53] Nevertheless, in order for an arrangement to be original, one factor that should be considered is whether the arrangement was standard convention. For example, in *Feist* the alphabetization of the white pages was a standard convention. [54]

a. The Physical Arrangement of the Database

{32} A question that should be asked is are we talking about the physical arrangement of the database, or are we talking about the logical arrangement? [55] This question need not be asked when the *Feist*

analysis is applied to a factual compilation in book form. The arrangement at issue in books is the physical arrangement; that is, how the different sections within the compilation are arranged, and how the material in each section is arranged. The introduction of factual compilations into a database adds the new and substantial factor of logical arrangements. The logical arrangement will be addressed in the material following.

{33} The physical arrangement refers to how the data is physically arranged on the storage device. When analyzing the physical arrangement of a database, one would be looking at directories, files, and how different structures are arranged within the files. A portion of the physical arrangement may be determined by a particular database management system that could have been used in creating the database. The goal of a good database management system seems to be to take care of the physical arrangement as much as possible without requiring extensive user direction so that the database creator can concentrate at the logical level. [56] From this one may infer that the logical arrangement may be the level at which a database arrangement was truly designed. [57] Usually, a database management system will provide a database creator with a number of structures and arrangements that may be used in constructing a database. Database developers typically simply use these basic building blocks in creating a database. [58] The content to be considered, in determining if the physical arrangement was original in a *Feistian* sense, should not include that part of the arrangement that was determined by the database management system being used because the author did not contribute this to the factual compilation. In short, the physical arrangement should only be considered in a *Feist* inquiry if the database creator designed that physical arrangement.

{34} Another problem with the physical arrangement is that it is often highly functional in nature. Developers often select the particular physical arrangement to maximize the speed and efficiency of the database interface program and its access to the database. For example, if a particular database was used mostly by users inputting a specific search and/or a specific arrangement, the database creator may maximize efficiency of the database package by arranging the database to optimize performance with such a search. An arrangement made based on efficiency may be closely tied to the idea or process calling for that arrangement, and may, therefore, not deserve copyright protection. [59] Even if these arrangements do deserve copyright protection, they should be scrutinized under the stricter tests for functional subject matter. [60] More likely than not these arrangements will contain sufficient authorship to warrant some copyright protection. [61]

{35} Back to the SmallTown Engineers Database, assume that the database comes on CD-ROM. Further assume that the database consists of four data files: ee.dat, se.dat, me.dat, and ce.dat. These four files reflect the four tables shown in Table 1, namely the following tables: electrical (ee), software (se), mechanical (me), and chemical (ce). If the records within each file were in the order as shown in Table 1, then Table 1 would be a representation of the physical arrangement of the database.

{36} In applying a *Feistian* arrangement inquiry to this context, assume that the records within each table are not arranged in any purposeful order. Most likely, they ended up in that order simply because of the order in which they were entered into the database. The arrangement of the records into electrical engineers, software engineers, mechanical engineers, and chemical engineers would probably not pass the originality requirement. These groupings are much like the groupings in *BellSouth* and therefore, most likely, are not copyrightable.

b. The Logical Arrangement of the Database

{37} In considering whether the arrangement of a database is sufficiently original to merit copyright protection, a court may consider the logical arrangement of the database. The logical arrangement refers to what a person accessing the database "sees". The computer program used to access the database allows a user to see and/or create logical arrangement(s). The logical arrangement is akin to a "simulated arrangement" in that the data is not actually physically arranged as it appears in a view of the logical arrangement. Of course, there may be some similarity between the physical arrangement and the logical arrangement.

{38} The logical arrangement of an entire database may be quite complex. As one delves into the areas of relational and object-oriented databases, the logical arrangement becomes painfully complex. Therefore, for this paper, the simple conceptual model of a database as outlined in Figures 1-3 and Table 1 will be used. [\[62\]](#)

{39} As stated, the physical arrangement and logical arrangement may be related, but they may also be unrelated. Take again our example of the SmallTown Engineers Database. Assume that the database includes twelve files, with each file corresponding to an engineer. In this case, the arrangement shown in Table 1 would be a logical arrangement because the physical arrangement is not the same as shown in Table 1. The computer program may still allow the user to view the database conceptually as shown in Table 1.

{40} A user could also create a logical arrangement as the result of a search. Assume the user wishes to search for all the engineers who like Star Trek. Once the search is performed, the computer program will display a portion of the results (or all of the results depending on the number of results) on the screen. To the user, it appears that there is an arrangement of engineers who like Star Trek. To be sure, there is, but this arrangement is a logical arrangement that was accomplished "on the fly" by the computer program. The actual physical arrangement has no relation to the logical arrangement of trekkies just accomplished. The selection of the data in the database enables a database interface program to accomplish various logical arrangements resulting from searches. In this particular case, the fact that the database creator included the selection of the Star Trek field facilitated a computer program being able to accomplish a logical arrangement based on the data. The selection of data determines the logical arrangements that may be accomplished by the search software.

{41} Logical arrangements that are accomplished by user inputs and the search software are arguably achieved through procedures or processes. [\[63\]](#) In no case does copyright protection extend to procedures or processes. [\[64\]](#) If a copyright were granted to a compilation based on such a logical arrangement, it may be effectively prohibiting the process or procedure to accomplish such an arrangement.

{42} In addition, these logical arrangements accomplished by a computer under user direction are very mechanical in nature. In fact, a typical search results type of logical arrangement accomplished by a computer program is the quintessential mechanical, garden variety arrangement of which there is no copyright protection. This points to the conclusion that originality should not be found in the typical search results logical arrangement accomplished by a computer program at the request of a user.

{43} The idea/expression dichotomy [\[65\]](#) and the doctrine of merger [\[66\]](#) also cause concern for extending protection to these search results logical arrangements. The merger doctrine is a direct outgrowth of the idea/expression dichotomy and stands for the rule that when there is only one or a limited number of ways to express an idea, the expression "merges" with the idea and becomes unprotectible. [\[67\]](#) Often this arises in cases involving highly functional and utilitarian works. With a

computer program achieving a logical arrangement by executing a certain algorithm, it seems highly likely that there are only a limited number of ways to express that particular idea of arranging, and therefore, the expression of that idea merges with the idea and becomes unprotectible.

{44} Some courts prefer to find a "thin copyright," rather than a merger. [68] A thin copyright protects against close copying only, but merger may deny protection altogether. [69] *Feist* states that a copyright in a factual compilation is thin. [70] However, the court was considering the arrangement of a hardcopy telephone directory: a physical arrangement. Logical arrangements achieved through a computer program would seem to pull a court's thinking more toward the merger doctrine because they are enhanced and viewable through a computer program and are sometimes determined and achieved through an idea or process for sorting.

{45} Logical arrangements inherent in the arrangement of the database, not the search results types of logical arrangements, should be considered in a *Feist* arrangement analysis because the database creator authored this arrangement. With modern database management software, the logical arrangement was probably the primary focus of the database creator.

{46} The search results logical arrangements are a different story. These logical arrangements are facilitated by the selection of data made by the database creator. Therefore, a search results logical arrangement should be protected but only by looking at the selection of data provided in the data records that facilitate a computer program in portraying and/or accomplishing the logical arrangement. In other words, because these logical arrangements are determined by the selection of fields in a record, the true point of originality lies in the selection of fields, not in any logical arrangements that may be accomplished on these fields. A program cannot accomplish a search results logical arrangement without a set of data to perform some algorithm on. By looking at the selection to protect these types of logical arrangements, the *Feist* analysis is kept where it should be, on the static data stored on a storage medium: on the factual compilation. Looking at user-determined search-results logical arrangements pulls us into the computer program and into the user interface, and therefore, into differing copyright standards. [71]

{47} The logical arrangement is shown to a user through a user interface. A user interface displaying results to a user is trying to most effectively convey to the user the desired information in such a way as to minimize confusion and maximize utility. A software developer working on a database interface program, when designing the displaying format, is worrying about how "best" to show the results to a user in a utilitarian sense. A software developer is probably not at all, or at least only slightly, concerned with achieving a form of creative expression. The developer is trying to create a user interface that comports with the standards of the industry so a user will easily navigate and use the software.

{48} If we allowed the computer program and its functionality into the *Feist* analysis, two nearly identical databases may be treated differently depending on the database interface program. For example, if one database interface program allowed only limited searching and arranging, a court, if looking at the logical arrangement as created by the program, may find the database not copyrightable because of its lack of original expression in selection and arrangement. However, if another database interface program were used to access the exact same database, and this program was loaded with functionality in allowing a user many different logical arrangements, a court may find, if looking at the logical arrangements through and/or created by the program, that the database was copyrightable because there was original expression in the arrangement. Thus, the exact same database could, if courts do not treat the problem carefully, find different protection under the copyright laws because of the interface software. This should not happen because the database is data; it is a compilation of facts. To

determine the copyrightability of a compilation of facts one looks to the selection and arrangement of the compilation. One does not look at how or if a computer program, which is copyrightable, manipulates this data to determine the copyrightability of the database.

{49} In sum, a court should look at what the database creator designed. If the database creator designed the arrangement as a logical arrangement, then that should be the focus. If the database creator were at the lower level physical arrangement, then that should be the focus. The question is not so much whether it was a physical or logical arrangement, but the question is what did the database creator author. However, court's should keep in mind these distinctions so as not to be mindlessly led through all sorts of arrangements that the database creator had nothing to do with.

{50} The logical arrangement of a database was analyzed by the trial court in *Warren Publishing v. Microdos* [72] when the court was looking at the question of infringement. Warren Publishing published annually a printed directory known as the Factbook providing information on cable television systems throughout the country. The defendant, Microdos, markets a software package called Cable Access. Cable Access was like the Factbook in that it provided detailed information on the cable television industry. [73]

{51} The trial court pointed out that the Cable Access software package came pre-sorted by state and city. [74] This is most likely referring to the physical arrangement because of term "pre-sorted" used by the court. The trial court also pointed out that a user may rearrange the data in a format of its choosing: [75] the logical arrangement. This is a search results type of logical arrangement. Warren claimed that the Cable Access software infringed its copyright of the Factbook. [76] Microdos did not contest that Warren had a valid copyright in the Factbook, but it contested that it had copied any copyrightable portion of the Factbook. [77]

{52} In articulating the test for copyright infringement in this case, the trial court stated that a finding of infringement would be supported if there was substantial similarity in the selection and arrangement between the "Cable Access software package" and the Fact book. [78] A much better statement would have been comparing the database (without considering the software to access and use it) and the Factbook because in such a statement the court would have been much clearer on what was truly at issue: factual compilations.

{53} The trial court found that Warren's selection of communities was sufficiently creative to be copyrightable. [79] > The trial court seemed to correctly frame the issue as to the selection by stating that "infringement will turn on whether or not Microdos's selection of communities is substantially similar to Warren's." [80] The trial court found that there was infringement of the selection and accordingly granted Warren's motion for summary judgment on this issue. [81] In the interests of providing a complete record, the trial court went on to analyze the coordination and arrangement of the Factbook.

{54} The court noted that the Factbook's coordination and arrangement of its data fields were copyrightable, but held that there was no infringement on this issue. [82] The court, in pointing out the reasoning for this holding, compared the visual differences between the Factbook and the Cable Access user interface. [83] The trial court also said that the differences would be even greater because a user can specific options to only display a limited record rather than a full record. [84] The user interface and the functionality of the computer program were clearly factors, if not the main inquiry, in the trial court's decision regarding the arrangement. This is unfortunate because the *Feist* analysis was inappropriately applied to works that are not factual compilations.

{55} The eleventh circuit, on appeal, did not have the issue of arrangement of the data fields before it. The appellate court affirmed the trial court's decision and did not analyze the arrangement of the data fields, as the trial court did. The appellate court should have clearly distinguished between the computer program and the database it was accessing, [85] and further should have been careful to only apply the *Feist* analysis to the factual compilation only: the database.

C. The Vulnerability of Databases

{56} *Feist* makes clear that unless a factual compilation is copyrightable, it may be copied at will. [86] Moreover, even those factual compilations exhibiting the minimal degree of creativity required to be copyrightable are vulnerable to being copied if the originality of the author can be divorced from the rest of the work. To illustrate, in the SmallTown Engineers database, it would be easy for a user to copy all the data except for the Star Trek field, and thus, divorce the originality (the Star Trek field) from the rest of the work. With databases, this means that if the selection is not creative, then the database may be copied and rearranged. Even if the selection is creative, a second comer may copy a large portion of the database if the fields within the records that make the database original can be left behind. These statements, in combination with the ease with which electronic information can be copied and stored, create very little protection for databases of facts.

{57} For example, in our hypothetical SmallTown Engineers Database, the only copyrightable expression seemed to be the Star Trek selection of data included in each record. If a copier came along and copied the whole database to sell on his own CD-ROM, he would likely be found infringing of a copyright in the unique selection of data--the Star Trek information. However, if the copier is aware of the copyright standards for factual compilations, he can easily divorce the copyrightable portion from the noncopyrightable portion. For example, he could write a little program that sequentially goes through and extracts all the data from the database except the information about Star Trek. If done in this way, he has not infringed any copyrights by copying most of the directory because he avoided copying the copyrightable expression.

1. Selection

{58} The value of a database may depend on how complete the database is. Specifically, users may find more value in a database where they can decide what data they want to extract or see: where the user determines his or her own selection. This selection is made by the user through the functionality of the search software. To facilitate this, database creators may have an incentive to include all the data available for any given area. [87] This may create a problem for database creators: they may lose commercial value if they are not complete, but the less complete they are the more creative a selection they have shown. This situation may present a difficult choice to database creators of taking more commercial value for less copyright protection, or vice versa. [88]

2. Arrangement

{59} The coordination or arrangement criterion of the originality requirement may be a stumbling block for information compilers. A database, particularly a static database (e.g., one sold on CD-ROM), is likely to have some arrangement to facilitate rapid searching and/or access. However, some databases, particularly dynamic databases (e.g., on-line databases that are constantly being updated with new

material), may not have any purposeful physical arrangement. [89] For those databases lacking any purposeful arrangement, combined with including as much information as possible so that the selection is minimal or nonexistent, there seems to be little chance for any copyright protection.

3. Is This a Bad Thing?

{60} We should ask ourselves if such little protection for databases of facts is a bad thing? This effectively puts those uncopyrightable facts, selections, and arrangements contained in these databases into the public domain like no other medium ever could. This helps ensure that people have information and can disseminate it as much as is possible under the copyright laws. This free flow of information is highly desirable because it helps educate society; it puts the information "out there" so that it is available to everyone.

{61} Perhaps this thin protection may result in being a disincentive to those in the database industry. One may imagine countless hours of tedious data entry in creating a database, [90] only to have it all copied in a few short moments on a computer. Then a competitor avoids all the costs of developing the database, but gets the rewards of the original author's labors. This may be true in some cases. However, this vision of an original author spending countless hours amassing data by hand may not be as accurate as one would think. For example, there are probably database "creators" out there who electronically accumulate the data to be put in a database. As technology continues to change society, original entry of data will probably be in electronic form anyway. Thus, the first author of a type of database may have acquired all the data through the use of a few routines written to gather data from around the internet and place it in a few files in certain formats, much like a second database creator may do.

{62} There is still merit in the argument that work may be bypassed by a "free rider" riding on the works of an original database creator. However, *Feist* pointed out that the "primary objective of copyright is not to reward the labor of authors, but 'to promote the Progress of Science and useful Arts.'" [91] Therefore, the issue should be not whether someone's hard work is used by another in a seemingly unfair way, but should be whether science and the arts would be better off by extending more protection or leaving the standard where *Feist* left it. In other words, if the benefit to the public of increased access to information (brought about by competitors using the uncopyrightable material from other's works) outweighs the loss of some factual compilations because of authors lack of protection, *Feist* should be left where it is.

a. The Free Rider Problem

{63} To be fair to those in the database market, the free rider problem should be addressed. The free rider comes later and benefits from much of the work of the first database creator by simply copying the uncopyrightable facts in the database and compiling another competing database. Even if the first database met the originality requirement, a second database creator may still copy the underlying facts without incurring any copyright infringement liability. [92] The free rider has avoided a substantial cost that the original database creator incurred. [93] Although this free riding does seem to avoid economic waste, [94] allowing such behavior in the competitive market does not comport with typical ideas of fairness. *Feist* allows free riding to take place. [95] Such a policy may discourage the creation of databases. [96]

{64} Perhaps the *ProCD* [97] case is an exemplary example of free riding. ProCD "compiled information from more than 3,000 telephone directories into a computer database" and sold it. [98] Zeidenberg bought a copy of the database, extracted the data from the database, and sold information from the database on the Internet. The database of ProCD cost more than \$10 million to compile. [99] The database of Zeidenberg apparently cost him about \$150. [100] What a bargain. Justice was seemingly served, though, when the seventh circuit held Zeidenberg bound to the shrinkwrap license. [101]

IV. Attempts to Fix the Problem

{65} To help avoid this problem of easily misappropriating data from a database, legislation has been introduced on several fronts that would establish a *sui generis* right for the protection of databases. Three pieces of legislation will be discussed. The focus will be on H.R. 3531, a database protection bill that was introduced last year by Representative Moorhead of California. However, the language between the three pieces of legislation is so similar that a critique of one will very likely be a critique of the others as well. Accordingly, as criticisms are mentioned relative to one particular piece of legislation, realize that the criticism applies to the others also. Generally speaking, although meaning well, each piece of legislation introduced seems to use a meat cleaver where it should have been using a much more precise and discriminating scalpal.

A. The European Directive

{66} Before a critical look is given to H.R. 3531, summarizing the background in database legislation would be useful. In March 1996, the European Union ("EU") adopted the European Directive on the legal protection of databases. [102] "The European Directive [103] instructs the fifteen member nations of the European Union to harmonize their laws to a uniform standard of copyright protection for databases." [104] The Directive goes beyond copyright protection; the directive protects the sweat-of-the-brow, the doctrine rejected in *Feist*. [105] It is very likely that non-EU databases (e.g., a U.S. database) will not receive any of this *sui generis* protection unless their country extends equivalent protection to EU databases. [106] The directive provides protection to database creators only if they are located in an EU member state. "The E.C.'s reciprocity requirement thus puts considerable pressure on the U.S. to enact *sui generis* compilation protection." [107]

{67} The European Directive defines a "database" as "a collection of words, data or other independent materials arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means." [108] "The most controversial aspect of the Database Directive is the so-called *sui generis* right, which prevents unauthorized extraction or re-use of the data that comprises the database." [109] The Database Directive defines Extraction as "the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form," [110] and the Directive defines re-use as "any form of making available to the public all or a substantial part of the contents of a database by on-line or other forms of transmission." [111] The European Directive provides for a term of 15 years, renewable upon a showing of new investments in the database. [112]

B. The Database Treaty Considered at WIPO

{68} In December 1996, the World Intellectual Property Organization [113] ("WIPO") considered three treaties designed to improve copyright protection in the digital age. Two treaties were approved, but one treaty, on database protection, was tabled. [114] Bruce Lehman told reporters that the draft treaty on database protection had been "dropped out" of the current WIPO deliberations following objections "from almost all countries of the world except Europe." [115] If a database treaty were adopted by WIPO, the United States would likely be under additional pressure to enact *sui generis* database protection. [116]

{69} The proposed treaty would have obligated countries to protect "any database that represents a substantial investment in the collection, assembly, verification, organization or presentation of the contents of the database." [117] A database is defined as a "collection of independent works, data or other material arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means." [118] The WIPO Draft Treaty would have given a database maker the right to "authorize or prohibit the extraction or utilization of its contents." [119] Extraction is defined as "the permanent or temporary transfer of all or a substantial part of the contents of a database to another medium by any means or in any form." [120] Utilization "means the making available to the public all or a substantial part of the contents of a database by any means." [121] A substantial part "means any portion of the database, including an accumulation of small portions, that is of qualitative significance to the value of the database." [122] The treaty also provided that any substantial change, qualitatively or quantitatively, which is a new substantial investment will qualify the database for a new term of protection. [123]

{70} According to Lehman, the principal reason for tabling the database proposal was because there was not enough time during the conference to work out the details. He acknowledged strong opposition to the treaty among the "hundreds" of U.S. industry lobbyists who gathered in Geneva for the conference. The Clinton administration "backed away from endorsing it" after the presidents of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine expressed their grave reservations that it "would seriously undermine the ability of researchers and educators to access and use scientific data, and would have a deleterious longterm impact on our nation's research capabilities." [124] Lehman suggested that the failure of the database treaty "gives us much more time and freedom" to evaluate the issue. However, he acknowledged the European Directive and its reciprocity, and that this could force movement on database legislation in this country. [125]

{71} "This treaty represents 'the end of the public domain,' according to American University Law Professor Peter Jaszi." [126] James Love, director of the Consumer Project on Technology, part of the Center for Study of Responsive Law in Washington gave several examples of what could happen under the proposed treaty. Love stated that stock quotes, "now disseminated freely, would be owned by stock exchanges and subject to license." [127] He gave another example involving sports statistics and newspapers in stating that "newspapers would need to get permission from professional sports leagues to print sports statistics." [128] John Browning reported that internet companies were concerned that "under the vague terms of the treaty, even Web pages could be considered databases, which might restrict browsing." [129]

C. H.R. 3531

{72} Representative Moorhead introduced H.R. 3531, the "Database Investment and Intellectual Property Antipiracy Act of 1996" in May of 1996. H.R. 3531, [130] in simple terms, protects the sweat of the brow. H.R. 3531 closely parallels the proposed treaty at WIPO. This bill was not passed and has not been reintroduced, [131] yet. [132] "Internet access providers and the scientific and academic

communities applaud the failure of the 104th Congress to pass" H.R. 3531. [133] As one might imagine, players in the information industry are hoping that a bill like H.R. 3531 is passed. [134] The information industry has been calling for database protection for a long time. [135] When introducing the bill, Representative Moorhead stated that the bill would "prohibit the misappropriation of valuable commercial databases by unscrupulous competitors who grab data collected by others, repackage it, and market a product that threatens competitive injury to the original database." [136] This is what should be done. However, H.R. 3531 seems to go a lot farther than merely attempting to accomplish these worthy goals. [137] A closer look will be taken at several provisions of H.R. 3531.

1. Statutory Authority

{73} H.R. 3531 was an exercise of the plenary power granted to Congress under the commerce clause. Although a full analysis of the statutory authority used as a basis for H.R. 3531 and the bill's constitutionality is beyond the scope of this paper, it should be noted that in *Feist*, Justice O'Connor stated that originality was a constitutional requirement. This being the case, it is questionable whether Congress can do an "end-run" around the constitutional constraints on copyright protection and enact H.R. 3531. [138] Professor Ginsburg thinks that Congress does have the authority to enact a carefully crafted statute aimed at protecting factual compilations. [139]

2. Critique of H.R. 3531

{74} H.R. 3531 defined a database as a collection of anything arranged systematically. [140] A database is subject to the bill if "it is the result of a qualitatively or quantitatively substantial investment of human, technical, financial or other resources." [141] The definition of a database to be protected by the bill could not be more broad. [142] every database creator would assert that their database was the result of a qualitative or quantitative investment. It seems that as long as data on a storage medium has some sort of structure and that the data just didn't appear there, the data would be subject to H.R. 3531. Arguably user interfaces could be protected under a bill like H.R. 3531, as well as menu-command structures. [143]

{75} Although the bill's definition of database should be criticized because of its broad scope, the bill should be commended for its provision touching on computer programs. "Computer programs are not subject to this Act" including those programs used in combination with a database. [144] Hopefully this was meant to make it clear that in analyzing databases under this Act, a court was to look only at the database and not at the computer program used to interface with the database.

{76} The bill states that no person shall "extract, use or reuse all or a substantial part, qualitatively or quantitatively, of the contents of a database subject to this Act" in a way that "conflicts with the database owner's normal exploitation" or "adversely affects the actual or potential market for the database." [145] If a person took important pieces of information from a database, but only a small portion of the database (e.g. less than 5%), would that be an extraction of a qualitative substantial part? It seems that almost any information worth extracting and recompiling would be a qualitative substantial part. One might think that if a good definition could be found for a qualitative substantial part, a compiler of facts might know a line over which he or she should not go. Perhaps as long as the extraction of facts was insubstantial and not qualitatively substantial a compiler would be safe in gathering a few facts from a database. Perhaps not.

{77} H.R. 3531 also prohibits a person from repeatedly or systematically extracting "insubstantial parts, qualitatively or quantitatively" in a way that "cumulatively conflicts with the database owner's normal exploitation" or "adversely affects the actual or potential market." [146] It is doubtful that this provision could not be used to nab almost anyone gathering facts from the database who does not do it as a one-time extraction. Section 5(a) of the bill states that, subject to the provision just described, a user is not prohibited from using insubstantial parts of the contents of a database, for "any purposes whatsoever." [147] However, in light of the broad language of § 4(a)(2), the effect of § 5, if any, seems to be questionable.

{78} The bill provides that acts that conflict with the normal exploitation of the database or that adversely affect the actual or potential market for the database include the use of contents of a database in a product or service where the customers might otherwise be expected to buy the database. [148] Here, it seems that as long as the contents used are not readily available off the internet, or somewhere else as accessible, for free, it could be argued that the person might have been expected to buy the database.

{79} The duration of protection for a database subject to this act, as a practical matter, seems to be forever. [149] The bill limits the protection to 25 years, but successive 25 year terms may be added to the duration of protection of dynamic databases if the changes are of "commercial significance, qualitatively, or quantitatively." [150] Those databases that change over time (e.g., online databases) probably have changes rolled into them at least every few years and probably every few days. Arguably there are no changes made to a database that are commercially insignificant, or at least not any that could be proven without some drawn out litigation.

{80} The long term allowed for protection under this Act may begin to look like copyright protection. The more the protection looks like copyright protection, the more likely a court may strike down the law as unconstitutional. [151]

{81} H.R. 3531 seems to put a nice lock and key around facts contained in databases. One might say that people could still get the facts from books. However, as technology moves forward, more and more factual information will be available on electronic media and less and less in hardcopy format. In addition, electronically stored information is much more accessible to the public than hardcopy, and so it is highly likely that the electronic sources will be used much more than the hardcopy sources will. This is especially true in light of the searching capabilities provided by the computer program providing access to these databases. This great benefit to society of having free access to information quickly, and in searchable form, may be greatly hampered by a bill like H.R. 3531. Another criticism of H.R. 3531 is that it does not recognize fair use. [152]

{82} Because of the many policies set forth in *Feist*, a statute directed towards protecting the information industry against free riders should be carefully tailored so as to not lock up factual information. The free flow of information may be substantially hampered by H.R. 3531, which may serve to put a serious plug in the information superhighway. If a bill like H.R. 3531 were enacted, the effects on the free flow of information may be worse than the free-rider problem. [153]

D. Some Suggestions for the Next Database Protection Bill.

{83} To balance the interests, a statute should be enacted that clearly only targets those free riders copying facts from databases, repackaging them and then reselling them in competition with the first compiler. Several key provisions should be included in the next database protection bill introduced.

Using H.R. 3531 as a baseline, several changes to H.R. 3531 should be made that would balance the interests of the general public with the commercial database creator.

{84} First, the definition of database should be considerably narrowed. A database should only be subject to the Act if it comprises an electronic compilation of uncopyrightable facts. Furthermore, the Act should only apply to those portions of the database that contain uncopyrightable facts:

{85} Second, the Act should narrowly target those commercial competitors who free ride off the sweat of the first compiler. **[154]** This may be done by attempting to pick out those copiers who are offering substantially the same product as the first compiler and who are selling that product. It also seems desirable to try and draw a few clear lines here, so that database creators don't have to litigate every time they want to find out if someone has been misappropriating data. The Act could target those second compilers who offer products including a database (or access to a database) for sale, where 25% or more of a second compiler's database was taken from a first compiler's database, or where the second compiler took 50% or more of a first compiler's database. This type of provision would focus on commercial competitors, and those outside this class would be free to use the information without worry of being liable for misappropriation of data.

{86} Third, the term of the protection should be greatly reduced from the 25 year term in H.R. 3531. **[155]** The many policies set forth in *Feist* need to be protected: facts need to be freely accessible to people. By focusing on the commercial competitors, these suggestions have helped address that problem somewhat. Nonetheless, even commercial competitors should have access to facts in electronic form after some short period of time. **[156]** The Act should provide for a 4-year term of protection. The term should not be renewable. For those databases that add new facts (whether they are wholly new or just a modification of the old facts), the 4-year term starts from the entry of the new data into the database. This may complicate matters for the second compiler who wishes to wait out the 4 years and then copy the database because some of the data in the database will undoubtedly have a term extending beyond the original data.

{87} Finally, a compulsory licensing provision should be included for those situations where a first compiler had exclusive access to facts, or where the first compiler had an unfair advantage in having access to facts. This licensing provision should be able to remedy situations like that of *Feist* where a compiler had a clear advantage in its access to the facts and was unwilling to license the listings to a second compiler.

V. Conclusion

{88} This article has conveyed a basic understanding of how a typical database can conceptually be viewed in understanding the selection and arrangement standard for factual compilations. Because a database is used with a computer program and a user interface, courts should be careful not to allow output from the computer program through its user interface to be considered in a *Feist* analysis. This can be done by "turning off the computer" before any selection and arrangement considerations are made.

{89} Although *Feist* serves several important policies of the constitution and copyright law, it does allow free riding to occur. In an attempt to fix the problem, H.R. 3531 was introduced last year. Unfortunately, this bill has many shortcomings that need to be fixed before it is reintroduced. Particularly, the bill should narrowly focus on commercial competitors for a limited time instead of

using language broad enough to cover almost anyone using a database for a long, long time. It is clear that pressure from the industry, plus international pressure stemming from the EC Directive, will force some new efforts for database protection, although constitutionality remains a major hurdle.

ENDNOTES

[1] See 1 Howard B. Abrams, *The Law of Copyright*, § 2.02[C][3][c][iv] at 2-26.34 (Release #6, 9/96) (giving example of database including copyrighted elements and stating that the copyright standard for factual compilations need not be reached in this case.)

[2] 17 U.S.C. § 102 (1996).

[3] 17 U.S.C. § 101 (1996).

[4] 17 U.S.C. § 103(b) (1996).

[5] *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 499 U.S. 340, 358-59 (1991).

[6] *Id.* at 343.

[7] *Id.* at 344.

[8] *Id.* at 363.

[9] *Id.* at 354.

[10] *Id.* at 347.

[11] *Id.* at 350.

[12] *Id.* at 362-63.

[13] 999 F.2d at 1441.

[14] *Id.* at 1438.

[15] *Id.*

[16] *Id.* at 1439.

[17] *Id.*

[18] *Id.* at 1441. The court also held that the merger doctrine applied and that "the arrangement . . . 'merged' with the idea of a business directory, and thus is uncopyrightable." *Id.* at 1442.

[19] 945 F.2d at 516 (2d Cir. 1991).

[20] *Id.* at 515.

[21] *Id.* at 513.

[22] The court's reference to "mechanical" regarding arrangement will be of particular interest when we

start talking about databases. If the arrangement of a database was accomplished by a computer program, such as a sorting routine, the arrangement is de facto mechanical, otherwise the arrangement could not have been accomplished by a computer program.

[23] *Id.* at 514.

[24] *Id.* at 515. In the directory at issue, only three of the 260 categories and 1500 of the 9000 listing were duplicates.

[25] *Id.* at 516.

[26] 914 F. Supp. 665 (D. Mass. 1995).

[27] *Id.* at 669.

[28] *Id.* at 675.

[29] *Id.* at 669.

[30] *Id.* at 675.

[31] *Id.* at 669.

[32] *Id.*

[33] *Id.* at 677.

[34] *Id.*

[35] *Warren Publishing, Inc. v. Microdos Data Corp.*, 52 F.3d 950 (11th Cir. 1995), *vacated and reh'g en banc granted*, 67 F.3d 276 (11th Cir. 1995). Decided by 115 F.3d 1509 cert denied 118 S. Ct. 397 (1996)

[36] *Warren Publishing, Inc. v. Microdos Data Corp.*, 67 F.3d 276 (11th Cir. 1995)

[37] This is not to say that they would necessarily agree on a result, however.

[38] If you are a Windows user, this is done by double clicking on the program icon.

[39] There may be times when the line between the software and the database becomes a bit blurred. For example, the distinction becomes ambiguous when a computer program contains extensive data, or when a database contains little programs (different pieces of executable code). 1 William F. Patry, *Copyright Law and Practice* at 201 (1994). However, for this paper, it will be assumed that the line can clearly be drawn between the two.

[40] 17 U.S.C.A. § 101; H.R. Rep. No. 1476, 94th Cong., 2d Sess. 54 (1976). Computer programs have their own specific standards for copyrightability. The copyrightability of computer programs is outside the scope of this article. For a detailed analysis used by some courts see *Computer Assocs. v. Altai, Inc.*, 982 F.2d 693 (2d Cir. 1992) (the Second Circuit's abstraction-filtration-comparison test for non-literal aspects of computer programs) and *Gates Rubber v. Bando Chemical Indus.*, 9 F.3d 823 (10th Cir. 1993) (adopting abstraction-filtration-comparison test). For more information on this topic see 1 Paul

Goldstein et al., Copyright § 2.15.2 at 2:188 (2d ed. 1996) (section devoted to computer programs).

[41] Warren Publishing, Inc. v. Microdos Data Corp., Civ. A. No. 1:90-CV-1654-JOF, 1992 WL 235745, (N.D. Ga), *aff'd*, 52 F.3d 950 (11th Cir.), *vacated and reh'g en banc granted*, 67 F.3d 276 (11th Cir. 1995).

[42] Lotus Dev. Corp. v. Borland Int'l, Inc., 49 F.3d 807 (1st Cir. 1995), *aff'd by an equally divided Court*, 116 S. Ct. 804 (1996); Goldstein, *supra* note 26, § 2.15.3.1 at 2:199-200 ("Despite their close connection, computer programs and the displays they generate are discrete objects of copyright protection.")

[43] Goldstein, *supra* note 26. A detailed analysis of the copyrightability of user interfaces and the standards to be applied thereto is outside the scope of this paper. For more information on this topic see Goldstein, *supra* note 26 § 2.15.3 beginning at 2:199 (section devoted to computer generated displays).

[44] Remember that this article assumes the data is factual and not copyrightable.

[45] Claude Delobel et al., Databases: From Relational to Object-Oriented Systems § 1.2.1 at 8 (1995) (stating that the physical level is "how the data is represented on the storage device used by the computer system . . . [d]efined in terms of files and records.")

[46] Some database literature refers to three different levels: the physical level, the conceptual level and the external level. Claude Delobel, *supra* note 31 § 1.2 at 8; Catherine M. Ricardo, Database Systems: Principles, Design, & Implementation, at 127-28 (1990). However, for this article referring to only two levels will be simpler and most helpful.

[47] If the results could all be displayed at once, the portion displayed would be all the results.

[48] *Feist*, 499 U.S. 340 (1991).

[49] *Id.*

[50] 1 Abrams, *supra* note 1, § 2.02[C][3][a][i] n.89 (1996) ("It is not clear what the difference is between the coordination of data and the arrangement of data."); 1 Patry, *supra* note 25, at 200 ("it is unclear what difference there is between" coordination and arrangement.)

[51] *See* 1 Patry, *supra* note 25 at 196 (1994) ("As a general principle, the greater the amount of material from which to select, coordinate, or arrange, the more likely it is that a compilation will be protectible.")

[52] 1 Patry, *supra* note 25 at 200.

[53] 1 Patry, *supra* note 25 at 202 ("Arrangement and location are, for the most part, meaningless, with retrieval frequently accomplished by means of a complementary computer program."); L. Ray Patterson, *Copyright Overextended: a Preliminary Inquiry into the Need for a Federal Statute of Unfair Competition*, 17 U. Dayton L. Rev. 385 (1992) (referring to the Westlaw and Lexis databases he says that "[w]ithin the database there is no coordination or arrangement, and the user makes the selection from a database that contains the universe for the topic").

[54] *Feist*, 499 U.S. at 362 (referring to the white pages as "garden variety" and "devoid of even the slightest trace of creativity.")

[55] See *supra* notes 31 and 32 and accompanying text. The physical arrangement refers to the arrangement at the physical level, the logical arrangement refers to the arrangement at the logical or conceptual level. Specifically, the term physical arrangement refers to the actual arrangement of the database on the storage device (i.e., how the files are arranged and how the data in each file is arranged). The term logical arrangement refers to the arrangement as seen by a user of the database.

[56] See Claude Delobel, *supra* note 31, § 1.3.2 at 11 (stating that a database management system (DBMS) "also provides a language for specifying the physical organization of the data. This physical organization involves describing the storage devices, the placing of data, the access methods and the file organization (inversion, chaining, B trees, and so on). System administrators . . . can choose from among these implementation options. Obviously, the ultimate aim is that the DBMS should take complete responsibility for physical organization, but existing systems are far from achieving this ideal.") A DBMS "is the software package used to interact with the database." *Id.* § 1.1 at 4.

[57] See Claude Delobel, *supra* note 31, § 1.2.2 at 9 (referring to the logical representation as the "main part of the DBMS.")

[58] See Rama Velpuri, *Oracle Backup & Recovery Handbook*, at 34 (1995) (referring to the Oracle DBMS and what structures it provides database creators to work with this book states that "[t]he data in the database is stored in Oracle blocks. An Oracle block is the smallest unit of physical space, and is a multiple of operating system blocks A contiguous collection of Oracle blocks is called an extent. This is a logical unit of space. A collection of extents is called a segment, which is the next level of logical storage unit.") "When an object (such as a table . . .) is created, an extent is allocated to the object's segment." *Id.* at 35.

[59] *Cf.* *Computer Assocs.*, 982 F.2d at 708 (with computer program design "the concept of efficiency is akin to deriving the most concise logical proof Thus, the more efficient a set of modules are, the more closely they approximate the idea or process embodied in that particular aspect of the program's structure.")

[60] See Goldstein, *supra* note 26, § 2.16.1.1, 2:212 n.26 ("The arrangement in the underlying database that can be protected as a compilation is properly characterized as a functional . . . work. . . . As a consequence, protectible subject matter should be measured according to the more stringent test employed for functional subject matter.")

[61] See Goldstein, *supra* note 26, § 2.16.1.1 at 2:211-12 (opining that most of the arrangements of databases made for increased efficiency "probably contain sufficient authorship to satisfy the originality requirement.")

[62] However, arguably the relational and object-oriented databases are themselves structured the same way as the conceptual model outlined. They simply add additional fields in each record for relationships.

[63] See Robert M. Mattison, *Understanding Database Management Systems: An Insider's Guide to Architectures, Products, and Design*, at 19 (1993) (stating that the computer program used with a database "manages the storage and retrieval of a set of data, organizes it logically, and provides the user with certain functionalities to guarantee that the data will be logically organized and consistently applied.")

[64] 17 U.S.C. § 102(b).

[65] Baker v. Selden, 101 U.S. 99 (1879).

[66] Morrissey v. Procter & Gamble Co., 379 F.2d 675 (1st Cir. 1967).

[67] *Id.*; Margreth Barrett, Intellectual Property Cases and Materials 388 (1995).

[68] Barrett, *supra* note 53 at 389.

[69] *Id.*

[70] *Feist* at 349.

[71] *See supra* notes 26-29 and accompanying text.

[72] Warren Publishing, Inc. v. Microdos Data Corp., Civ. A. No. 1:90-CV-1654-JOF, 1992 WL 235745, (N.D. Ga), *aff'd*, 52 F.3d 950 (11th Cir. 1995). Although the 11th circuit summarized this part of the lower court's opinion, the findings with respect to the data fields were not before the court for review.

[73] Warren I, at 1.

[74] *Id.* at 2.

[75] *Id.*

[76] *Id.* I would have used more specificity in detailing exactly what part of the software package was infringing, but because the trial court and the appellate court both treated the software package as one "thing," it appears both courts lumped the database, the database interface program and its user interface into one unit.

[77] *Id.*

[78] *Id.* at 3.

[79] *Id.*

[80] *Id.* at 5.

[81] *Id.* at 7.

[82] *Id.* at 10.

[83] *Id.* at 7-8. (Conducting a visual comparison of the two works and noting that the Factbook had a four-column format for the page and a single column for the individual community entries; later pointing out that Cable Access provided "in a page-wide format, general information under the 'systemname'; then in a three-column format provides subscriber information on the left, with cable system and equipment information in the center, and community information on the right").

[84] *Id.* at 10.

[85] *See ProCD, Inc. v. Zeidenberg*, 908 F. Supp. 640, 647 (W.D. Wis. 1996), *rev'd*, 86 F.3d 1447 (7th Cir. 1996) (distinguishing between the computer program and the database and stating the "[t]he

difference between these two elements is critical").

[86] Feist, at 359.

[87] Jane C. Ginsburg, *No "Sweat"? Copyright and Other Protection of Works of Information after Feist v. Rural Telephone*, 92 Colum. L. Rev. 338, 345 (1992) (noting that information compilers may eschew selectivity in favor of comprehensiveness because the comprehensiveness may be more attractive to the user); see also Patry, *supra* note 25, at 202 (stating that databases may get their value from their comprehensiveness--their very lack of selection).

[88] Abrams, *supra* note 1, § 2.02[C][3] at 2-26.30 (Release #6, 9/96) (stating that "this may create a difficult problem for suppliers of data . . . because their commercial value may depend upon their completeness, even though they may lose copyright protection precisely by being complete."); see also James E. Schat et al., *What's Mine Is Yours? The Dilemma of a Factual Compilation*, 17 U. Dayton L. Rev. 423, 433-34 (1992) (noting that often databases are comprehensive and that "their comprehensiveness may be used against them in determining whether the compiler exercised original "selection" in creating the work.)

[89] Ginsburg, *supra* note 73, at 345 (stating that "some compilations, particularly computerized databases, may lack any 'arrangement.'"); see also Schat, *supra* note 74, at 434 (noting that "it may be difficult to demonstrate any meaningful original arrangement of data contained within a computer database; the arrangement of data may tend to be important only from a computer science standpoint").

[90] Schat, *supra* note 74, at 433 (stating that databases "require the investment of vast amounts of time and effort in order to collect and assemble the data").

[91] Feist, at 349 citing Art I, § 8, cl 8.

[92] Goldstein, *supra* note 26 § 2.14, at 2:153 (concluding that a competitor would be free to copy the underlying data in a database meeting the *Feist* standard and would face no copyright liability).

[93] Schat, *supra* note 74, at 434 (noting that databases, "though difficult and expensive to create, may be extremely cheap to steal."); see also J.H. Reichman & Pamela Samuelson, *Intellectual Property Rights in Data?*, 50 Vand. L. Rev. 51, 54 (1997) (mentioning recent initiatives to protect databases and stating that the initiatives "aim to rescue database producers from the threat of market-destructive appropriations by free-riding competitors who contributed nothing to the costs of collecting or distributing the relevant data").

[94] Goldstein, *supra* note 26, § 8.1.1.2 at 8:61 ("*Feist* eliminates the economic waste of a rule that required competitors to reinvent the wheel, incurring the same expenses as the first copyright owner to produce an identical product").

[85] Ginsburg, *supra* note 73 at 349 (stating that "Justice O'Connor's opinion appears to enshrine a policy of free-riding in the Constitution").

[96] Goldstein, *supra* note 26, § 2.14.2 at 2:161 (asserting that some form of property right is necessary to stimulate the production of databases and stating that *Feist* "introduced a measure of instability into the country's increasingly important data collection industries."); see also J. H. Reichman and Pamela Samuelson, *Intellectual Property Rights in Data?*, 50 Vand. L. Rev. 51, 55 (1997), (stating that "[t]he risk of market failure inherent in this state of chronic under-protection tends to keep the production of

information goods at suboptimal levels").

[97] ProCD, Inc., 86 F.3d 1447.

[98] *Id.* at 1449.

[99] *Id.* at 1449.

[100] *Id.* at 1449-50 (stating that ProCD sold to the general public the database for around \$150 and noting that Zeidenberg bought a consumer package).

[101] *Id.* at 1447.

[102] Commission Directive 96/9/EC of March 11, 1996, on the Legal Protection of Databases, 1966 O.J. (L077) 20.

[103] For a detailed look at the evolution and legislative history of the European Database Directive see Reichman and Samuelson, *supra* note 82, starting at 76. This article states that the "legislative history illustrates how a modest, pro-competitive initial proposal for sui generis protection has been transformed into a virtually absolute monopoly by the backdoor lobbying efforts of publishers and by the coordinated efforts of U.S. and E.U. officials to propogate a protectionist strategy for the global information infrastructure." *Id.* at 75.

[104] Gorman & Ginsburg, Copyright for the Nineties Cases and Materials 158 (Supp. 1996).

[105] *Id.*

[106] *Id.*

[107] *Id.*

[108] European Directive, art. 1(2).

[109] Richard Raysman & Peter Brown, *European Database Protection*, N.Y.L.J., Feb. 13, 1996, at 3.

[110] European Directive, art. 7(2)(a).

[111] *Id.* at art. 7(2)(b).

[112] *Id.* at art. 7(1).

[113] For more information on WIPO, see the WIPO home page (visited April 22, 1997) <<http://www.wipo.int>>.

[114] BNA Patent, Trademark & Copyright Law Daily, WIPO: WIPO Delegates Agree on Two Copyright Treaties (Dec. 24, 1996).

[115] *Id.*

[116] Ron Reiling, *Intellectual Property Regimes for the Information Age: Policies of the United States, the European Union and the World Intellectual Property Organization*, 3 B.U. J. Sci. & Tech. L. 9 (1997) ("The new WIPO treaties are likely to result in changes to United States and European

Community law in the near future.").

[117] Draft Treaty on Intellectual Property in Respect of Databases, WIPO CRNR/DC/4, Art. 1.1 (Aug. 30, 1996) ("WIPO Draft Treaty").

[118] *Id.* at art. 2(i).

[119] *Id.* at art. 3(1).

[120] *Id.* at art. 2(ii).

[121] *Id.* at art. 2(vi).

[122] *Id.* at art. 2(v).

[123] *Id.* at art. 8(3).

[124] Pamela Samuelson, *Big Media Beaten Back*, *Wired*, Mar. 1997, at 61, 184.

[125] Samuelson, *supra* note 100.

[126] Joe Celko, *Random data and observations; SQL for Smarties; News Briefs; Column*, *DBMS*, Mar. 1997, at 18.

[127] Tani Freedman, *Global Copyright Treaties Slammed by US Lobby*, *Agence France Presse*, Dec. 5, 1996, International news section.

[128] *Id.*

[129] John Browning, *Africa 1 Hollywood 0*, *Wired*, Mar. 1997, at 61, 185.

[130] H.R. 3531, 104th Cong. (1996).

[131] As of May 5, 1997.

[132] Julius J. Marke, *Database Protection Acts and the 105th Congress*, *N.Y.L.J.*, March 18, 1997, ("It is believed that a bill similar to the database protection proposal (H.R. 3531) will be introduced in the 105th Congress."); Samuelson, *supra* note 110 at 184 ("Database-protection legislation is also likely to be introduced in the 105th Congress.").

[133] Marke, *supra* note 118 at 5.

[134] The Information Industry Association's home page (visited May 26, 1997) <<http://www.infoindustry.org>>.

[135] *Copyright Law Revision, Part 1: Hearings on H.R. 2223 Before the Subcomm. on Courts, Civil Liberties, and the Administration of Justice*, 94th Cong., 1st Sess. (1975) (statement Paul G. Zurkowski, President, Information Industry Association) (calling for the protection of databases).

[136] 142 Cong. Rec. E890 (daily ed. May 23, 1996) (statement of Rep. Moorhead).

[137] Reichman and Samuelson, *supra* note 82 at 55 (asserting that H.R. 3531 is "seriously flawed" and

that its implementation "would confer a far broader and stronger monopoly on database developers than is needed to avert market failure").

[138] Ginsburg, *supra* note 73 at 341 ("Federal law protection under a revised copyright statute, or under a Commerce Clause-based misappropriation statute, may prove difficult to reconcile with Feist's constitutionally derived endorsement of free-riding on previously gathered information").

[139] *Id.* (concluding "that Congress does retain the authority to award a carefully tailored property right to information compilers.") For a detailed analysis of Congress' power to enact legislation to protect factual compilations, *Id.* at 367 (section entitled "Congress' Constitutional Authority to Enact a Statute Protecting Against Misappropriation of Compiled Information").

[140] H.R. 3531 § 2.

[141] H.R. 3531 § 3(a).

[142] Pamela Samuelson, *Legal Protection for Database Contents*, 39 Comms. A.C.M. no. 12, pg. 17 (Dec. 1996) (stating that the "definition of 'database' in the U.S. legislation and proposal to WIPO is too broad").

[143] *Id.* (Arguing that user interfaces and menu command structures could be conceived as databases components under H.R. 3531.)

[144] H.R. 3531 § 3(d).

[145] *Id.* at § 4(a)(1).

[146] *Id.* at § 4(a)(2).

[147] *Id.* at § 5(a).

[148] *Id.* at § 4(b)(3).

[149] Marke, *supra* note 118 (stating that the Executive Board of the American Association of Law Libraries expressed concern that "the protection term of 25 years, renewable each time a significant change is made to the database, would provide protection that could last in perpetuity."). *But see NII Copyright Protection Act of 1995: Hearings on H.R. 2441 Before the Subcomm. on Courts and Intellectual Property of the House Comm. on the Judiciary*, 104th Cong., 2d Sess. 69 (1996) (testimony of Barbara A. Munder, Senior Vice President of the McGraw-Hill Companies on behalf of the Information Industry Association) (referring to the European Directive and stating that its term, 15 years, "is quite short in comparison to that afforded copyright owners").

[150] H.R. 3531 § 6.

[151] Ginsburg, *supra* note 73 (arguing that Congress retains power under the Commerce Clause to enact an anticopying statute that departs in significant ways from copyright law so as to not conflict with the Patent-Copyright Clause, and giving an example of such a departure being a duration of protection for ten years, a term significantly shorter than the typical copyright term).

[152] John Gibeaut, *Zapping Cyber Piracy: Copyright Protection in the Digital Age Has Turned into a Battle over Space Turf, Raising Constitutional Questions of Owners' Rights vs. Public Access*, 83

A.B.A. J. 60 (Feb. 1997). Later this article states "Lehman maintains, however, that database protection seeks only to compensate those who compile information for their labor and is not intended to reduce access to content. Fair use will not wither and die, he says, although those who have been getting data for free may have to get used to paying for it." *Id.*

[153] Patterson, *supra* note 39 at 410 (asserting that a real danger in enacting a trade-regulation statute for neo-copyright "is that corporate lobbyists would influence the proposed legislation to such an extent that the cure would be worse than the evil").

[154] Patterson, *supra* note 39 at 409 (contending that a trade-regulation statute for neo-copyright, which in his view includes databases of factual compilations, "should be only against competitors").

[155] H.R. 3531 § 6.

[156] Patterson, *supra* note 39 at 409 (contending that protection provided by a trade-regulation statute for neo-copyright, which in his view includes databases of factual compilations, "should be for a limited time (and far less than a century)."); Samuelson, *supra* note 128 (reviewing the *ProCD* case and stating that with *sui generis* protection for databases the *ProCD* "doesn't really need 15, let alone 25, years of protection for the listings in its directory" and further stating that "[a] far shorter term would undoubtedly provide ProCD with adequate lead time in which to recoup its investment").

Copyright 1998 by the Journal of Technology Law & Policy; and Wesley L. Austin.

