

## COPYRIGHT PROTECTION OF COMPUTER DATABASES AFTER *FEIST*\*

John F. Hayden \*\*

### INTRODUCTION

The increased sophistication of computer databases<sup>1</sup> has made them invaluable tools for a range of uses.<sup>2</sup> At the same time unprecedented advances in computer technology have helped to spur growth in the computer information industry. As computers have become more powerful, they have also become more affordable. As a result, systems previously found only in research labs are now standard fixtures in offices and homes. Increased accessibility has created vast numbers of users demanding additional products and services, and this has been a further push for expansion in the computer information industry as a whole.<sup>3</sup> Databases now assume vital importance in virtually every segment of the economy.<sup>4</sup> Their widespread use and critical importance

---

\* This Recent Development will be entered in the 1992 Nathan Burkan Memorial Competition.

\*\* Harvard Law School, Class of 1992.

1. For purposes of this Recent Development, a "database" is a collection of data stored in an electronic format, a "developer" is the person or party who creates the database, and a "database system" is a combination of a database and the software necessary to manipulate that database. "Database" and "computer database" are used synonymously.

2. Tools such as LEXIS and Westlaw and their impact on legal research exemplify databases' expanded role.

3. Revenues of the industry were expected to be about 9 billion dollars in 1990 and are expected to grow to 19.2 billion dollars by 1994. See Brief of the Information Industry Association and ADAPSO, The Computer Software and Services Industry Association, Inc., as *Amici Curiae* in support of neither party at 21 n.17, *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282 (1991) (No. 89-1909) (citing U.S. DEPARTMENT OF COMMERCE, 1990 INDUSTRIAL OUTLOOK, 29-3 (1990)) [hereinafter IIA/ADAPSO Brief].

4. A few examples of useful databases include: financial information databases that are "indispensable tools for investors, regulators and participants in all financial markets"; credit reporting systems that are the basis for hundreds of thousands of daily business decisions; demographic databases that "play a dominant role in marketing, fundraising and planning decisions"; bibliographic databases relied on by researchers and students; and economic and industrial databases that "underpin momentous decisions made every day at all levels of business and government." *Id.* at 8-9. See also Priscilla A. Walter, *Databases: Protecting An Asset; Avoiding a Liability*, 8 COMPUTER LAW, 10, 10 (Mar. 1991) (positing that "[v]irtually all businesses, and most individuals, own or use one or more forms of database regularly"); *id.* at 20 (listing a sampling of the variety of databases available).

necessitate some sort of legal protection against their misappropriation.<sup>5</sup>

Database development involves considerable effort and expense while database copying is quite easy and inexpensive. Effective development involves input by a variety of experts.<sup>6</sup> Initially, marketing experts must identify a market niche and potential customers. Next, these experts assess user characteristics and locate and evaluate potential sources of information. The developer then carefully selects data from available sources and combines his findings into a coherent whole.<sup>7</sup> When a comprehensive set of data is finally collected, other experts must put it into a computer-useable format. In addition to the costs of collection and organization, the developer may then incur further expense in obtaining access to proprietary data sources.<sup>8</sup> At the same time, with proper access, a database can be copied in a matter of minutes with very little effort. Databases thus need legal protection because they are simultaneously difficult to produce and easy to copy.

While the need to protect databases increases proportionately with the resources expended in developing them, their increased sophistication has hindered traditional methods of protection. Currently, contractual arrangements provide the primary source of protection. However, new methods of distribution are increasing the likelihood that parties not in contractual privity with a database's owner and not bound by these arrangements will have access to the database.<sup>9</sup> Therefore, non-contractual methods of protection must be employed to prevent these third parties from copying at will.<sup>10</sup>

---

5. This proliferation of databases also has negative implications. See generally Richard Lacayo, *Nowhere to Hide*, TIME, Nov. 11, 1991, at 34-38 (examining the privacy issues raised by databases). However, these implications are directed at regulating the use of databases rather than discouraging their protection.

6. "The successful development and distribution of a database often depends on the solution of complex technical and marketing problems [and] . . . also calls for a sophisticated knowledge of information science, . . . information seeking behavior, and of the details of storage and retrieval systems and computer programs." IIA/ADAPSO Brief, *supra* note 3, at 19.

7. See *id.* at 18. The brief gives an example of a demographic database which "may draw upon public sources such as census data, property tax records, and voter registration files, and upon a plethora of proprietary sources that contain data on the target group such as purchasing patterns, subscriptions to catalogs or periodicals, and the like." *Id.*

8. See *id.*

9. For example, on-line information "gateways" and portable compact discs will reduce developers' ability to control access to their databases. See *id.* at 21. Moreover, as their desire and ability to manipulate data improves, users of database services will demand complete access to entire databases rather than the current controllable ability to perform searches.

10. See generally Jane C. Ginsburg, *Creation and Commercial Value: Copyright Protection of Works of Information*, 90 COLUM. L. REV. 1863, 1918-22 (1990) (examining the difficulties of protecting factual compilations through contractual measures).

Though the proper method of intellectual property protection for many aspects of software has been the subject of considerable commentary,<sup>11</sup> databases currently fall within the domain of copyright.<sup>12</sup> Federal intellectual property protection is founded on a constitutional mandate to Congress 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."<sup>13</sup> The patent system protects the "discoveries" of "inventors" while the copyright system protects the "writings" of "authors."<sup>14</sup> Occasionally, a domain arises that defies neat

---

11. See, e.g., John M. Conley & Robert M. Bryan, *A Unifying Theory for the Litigation of Computer Software Copyright Cases*, 63 N.C. L. REV. 563 (1985); Duncan M. Davidson, *Protecting Computer Software: A Comprehensive Analysis*, 1983 ARIZ. ST. L.J. 611; Steven W. Lundberg et al., *The Copyright/Patent Interface: Why Utilitarian "Look and Feel" is Uncopyrightable Subject Matter*, 6 COMPUTER LAW. 5 (Jan. 1989); Peter S. Menell, *Tailoring Legal Protection for Computer Software*, 39 STAN. L. REV. 1329 (1987); Pamela Samuelson, *CONTU Revisited: The Case Against Copyright Protection for Computer Programs in Machine-Readable Form*, 1984 DUKE L.J. 663; Peter G. Spivack, *Does Form Follow Function? The Ideal/Expression Dichotomy in Copyright Protection of Computer Software*, 35 UCLA L. REV. 723 (1986); Jeffrey R. Benson, Note, *Copyright Protection for Computer Screen Displays*, 72 MINN. L. REV. 1123 (1988); Jorge Contreras et al., Recent Development, *NEC v. Intel: Breaking New Ground in the Law of Copyright*, 3 HARV. J.L. & TECH. 209 (1990); Brett N. Dorny & Michael K. Friedland, Recent Development, *Copyrighting "Look and Feel": Manufacturers Technology v. CAMS*, 3 HARV. J.L. & TECH. 195 (1990).

12. See generally Jack B. Hicks, Note, *Copyright and Computer Databases: Is Traditional Compilation Law Adequate?*, 65 TEX. L. REV. 993 (1987) (examining the application of copyright law to databases). While the software component of a database system potentially falls within the province of both copyright and patent, see Davidson, *supra* note 11, at 649-52, the database itself is unsuitable for patent protection. Patent protection is extended only to an invention that can be classified as a "process, machine, manufacture, or composition of matter." 35 U.S.C. § 101 (1988). See DONALD S. CHISUM, PATENTS § 1.01 (1990). That an informational work does not fit into any of these classifications is embodied in the rule that "printed matter" is not "within the statutory classes of patentable subject matter." *Id.* § 1.02[4]. For example, a new method of arranging directories is unpatentable. See *In re Russell*, 48 F.2d 668, 669 (C.C.P.A. 1931) (noting that "[t]he mere arrangement of printed matter on a sheet or sheets of paper, in book form or otherwise," does not merit patent protection). Thus, without patent protection, copyright is currently the only source of federal intellectual property protection for databases.

13. U.S. CONST. art. I, § 8, cl. 8.

14. Any disagreement over which of "Science" or "useful Arts" was to be promoted by protecting the "Writings" of "Authors" was resolved by the Supreme Court when it noted that copyright promotes both purposes. See *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282, 1290 (1991) ("The primary objective of copyright is . . . 'to promote the Progress of Science and useful Arts.'") (quoting U.S. CONST. art. I, § 8, cl. 8.). Prior to *Feist*, the belief, based on colonial usage and the structure of the clause itself, was that only "Science" was to be promoted. See 1 MELVILLE B. NIMMER & DAVID NIMMER, NIMMER ON COPYRIGHT, § 1.03 n.1 (1990). In this context, "Science" is "used in the sense of general knowledge rather than the modern sense of physical or biological science." *Id.* (quoting *Williams & Wilkins Co. v. United States*, 172 U.S.P.Q. (BNA) 670, 683 (Comm'r Ct. Cl. 1972), *rev'd on other grounds*, 487 F.2d 1345 (Ct. Cl. 1973), *aff'd by an equally divided court*, 420 U.S. 376 (1975)).

categorization as either science or art. Software is one such domain. Meanwhile, databases, which at their most fundamental level are simply traditional compilations or directories expressed in a modern electronic medium, are protected by copyright.<sup>15</sup>

Intellectual property law, though a logical and necessary<sup>16</sup> means of protecting databases, is of limited value in this respect because its reliance on copyright as the source of protection is inadequate. Databases are now protected by copyright as factual compilations, coverage of which has been troublesome even without the multiple complications presented by computer technology.<sup>17</sup> Protection of factual compilations is difficult because they lack the element of creative expression required in traditionally copyrightable works, such as novels, paintings, and songs. Any expression in a factual compilation is based on the compiler's selection and arrangement of data. However, this selection

---

15. "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." 17 U.S.C. § 101 (1988). The legislative history of the Copyright Act of 1976 specifically provides for protection of factual compilations and computer databases. "The term 'literary works' . . . includes catalogs, directories, and similar factual, reference, or instructional works and compilations of data. It also includes computer data bases . . ." H.R. REP. NO. 1476, 94th Cong., 2nd Sess. 54, reprinted in 1976 U.S.C.C.A.N. 5659, 5667. See also 1 NIMMER & NIMMER, *supra* note 14, § 2.04[C].

16. Traditional state law theories of unfair competition and misappropriation might be useful to protect databases absent contractual arrangements. However, the inconsistent protection offered by these remedies is less than ideal for an industry that is national in scope. There is also concern that these methods are preempted by § 301 of the Copyright Act which provides that "all legal or equitable rights that are equivalent to any of the exclusive rights within the general scope of copyright" are preempted by copyright. 17 U.S.C. § 301 (1988). See Henry Beck, *Copyright Protection for Compilations After Feist*, 8 COMPUTER LAW. 1, 7 n.3 (July 1991); IIA/ADAPSO Brief, *supra* note 3, at 21-22. See also *Financial Info., Inc. v. Moody's Investors Serv., Inc.*, 808 F.2d 204, 208-09 (2d Cir. 1986), cert. denied, 484 U.S. 820 (1987) (finding state law misappropriation protection preempted by § 301 of the Copyright Act even while finding no copyright in a compilation due to lack of authorship). A full discussion of the preemption issue is beyond the scope of this Recent Development. For analysis of this topic, see Henry D. Fetter, *Copyright Revision and the Preemption of State "Misappropriation" Law*, 25 BULL. COPYRIGHT SOC'Y 367 (1978); David E. Shipley & Jeffrey S. Hay, *Protecting Research: Copyright, Common-Law Alternatives, and Federal Preemption*, 63 N.C. L. REV. 125, 139-41 (1984).

17. See Richard L. Brown, *Copyright and Computer Databases: The Case of the Bibliographic Utility*, 11 RUTGERS COMPUTER & TECH. L.J. 17 (1985); Robert C. Denicola, *Copyright in Collections of Facts: A Theory for the Protection of Nonfiction Literary Works*, 81 COLUM. L. REV. 516 (1981); Ginsburg, *supra* note 10; Robert A. Gorman, *Fact or Fancy? The Implications for Copyright*, 29 J. COPYRIGHT SOC'Y 560 (1982); Celia D. Moore, *Ownership of Access Information: Exploring the Application of Copyright Law to Library Catalog Records*, 4 COMPUTER/L.J. 305 (1983); William Patry, *Copyright in Compilations of Facts (or Why the "White Pages" Are Not Copyrightable)*, 12 COMM. & LAW 37 (1990).

and arrangement is generally determined by functional considerations and is of doubtful creativity.<sup>18</sup> Furthermore, the Supreme Court substantially curtailed the breadth of protection for factual compilations in its decision in *Feist Publications, Inc. v. Rural Telephone Service Co.*,<sup>19</sup> which has exacerbated<sup>20</sup> the difficulties of applying copyright law to databases. First, it eliminated the "sweat of the brow" doctrine, which had provided the broadest protection to factual compilations. It also articulated a standard for protection which will be particularly difficult for databases to meet because it focuses on qualities whose proper application to databases is questionable. In addition, even if this standard is met, it provides virtually no protection of the data itself and thus fails to protect the most critical element of databases. A final limitation on copyright protection of databases results from the difficulty of using copyright to protect useful articles.

This Recent Development examines the current state of copyright protection for databases in light of *Feist* and some subsequent lower court decisions. In particular, it argues that copyright protection of databases is virtually nonexistent after *Feist* and that by failing to protect these vital resources, the copyright system discourages their development and dissemination and thus does not fulfill its constitutional mandate. As background, Section I explains the statutory basis for database protection. Next, Section II discusses the *Feist* opinion and its impact on compilation protection. Section III examines the standard set forth in

---

18. See 1 NIMMER & NIMMER, *supra* note 14, § 2.01[B] n.35 (citing numerous opinions for the proposition that there is insufficient creativity when expression is dictated by functional considerations). See also *infra* Section IV.

19. 111 S.Ct. 1282 (1991).

20. One commentator has noted that *Feist* "raises serious issues for vendors of financial information, proprietors of computer databases, and other publishers and compilers of essentially factual information. Beck, *supra* note 16, at 1. Moreover, Beck warns that the decision may lead to "a veritable 'gold rush' of information reorganization, reconstruction and redistribution." *Id.* at 2.

The database industry recognized these potential implications when it asked the Court to issue a narrow decision. See IIA/ADAPSO Brief, *supra* note 3, at 5. However, the Court ignored this request in rendering a sweeping opinion. The breadth of the opinion was noted by Register Oman in his address to the Atlanta meeting of the Patent, Trademark and Copyright Section of the American Bar Association:

The information industry submitted a number of *amicus* briefs in [*Feist*] almost begging the Court to issue a narrow decision not addressing databases. The Court declined the invitation in a big way . . . . What does this mean in practice? We're not sure yet. [T]he Court provided no guidance[,] . . . other than to state that the "selection" prong of the definition of compilation would be key. I anticipate a great deal of litigation in this area will be necessary to flesh out the issue.

*Feist* and expands upon this standard in light of other opinions. Section III also examines the particular impact of this standard on databases, and Section IV discusses the problems caused by the utilitarian nature of databases within the idea-expression dichotomy of traditional copyright law. Finally, Section V proposes a new system of *sui generis* protection for databases and factual compilations.

## I. SOURCE OF PROTECTION

Because they are generally composed of preexisting factual information,<sup>21</sup> databases are protected as factual compilations under the copyright statute.<sup>22</sup> Coverage of compilations<sup>23</sup> is explicitly limited to "material contributed by the author" and "does not imply any exclusive right in the preexisting material."<sup>24</sup> This means that the crucial issue in determining the scope of copyright in a compilation is establishing what was actually contributed by the author. Prior to *Feist*, courts had generally resolved this issue under one of two conflicting theories.<sup>25</sup> "Sweat of the brow" (or "industrious collection") opinions based expansive protection on the underlying notion "that copyright was a reward for the hard work that went into compiling facts."<sup>26</sup> Meanwhile, "selection and

---

21. A database composed of original materials would be protected under the category of compilations known as collective works. A collective work is a collection of "separate and independent works." 17 U.S.C. § 101 (1988). It receives separate protection for its individual components as well as protection for the overall work as a compilation. Because many, if not most, databases consist of factual or public domain materials that are individually unsuitable for copyright protection, they will not benefit from this additional layer of coverage. Even if a particular database consists of copyrightable materials, in many instances other parties will own the copyright in those materials and thus the developer will not benefit from their protection. Therefore, in discussing copyright protection of databases, this Recent Development treats them as "pure" compilations and assumes that they do not consist of individually copyrightable components.

22. Thus, for purposes of this Recent Development, "factual compilation" is a general term which includes databases. Meanwhile, "database" does not include all factual compilations.

23. See 17 U.S.C. § 101 (1988); *supra* note 15.

24. 17 U.S.C. § 103(b) (1988).

25. See Beck, *supra* note 16, at 3. This perceived split between the circuits was one basis for *Feist's* petition for *certiorari*, see Pet. for Cert. at 10, *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282 (1991) (No. 89-1909), and was the basis of the grant. 111 S.Ct. 40 (1990) (limiting the grant to the *Feist's* third question, the one which had addressed this split).

26. *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282, 1291 (1991). "Sweat of the brow" originated in an early Second Circuit opinion:

The right to copyright a book upon which one has expended labor in its preparation does not depend upon whether the materials which he has collected consist or not of matters which are *publici juris*, or whether such materials show literary skill.

arrangement" opinions limited protection to the originality that the compiler added in his selection and arrangement of facts.<sup>27</sup> In *Feist*, the Court eliminated "sweat of the brow" as a means of covering compilations.

## II. FEIST v. RURAL TELEPHONE

### A. Background

*Feist* involved a dispute between two publishers of telephone directories. A telephone directory is a classic example of a factual compilation under copyright law, making the decision in *Feist* directly applicable to databases. Rural Telephone Service Co., a telephone company in

---

or originality, either in thought or in language, or anything more than industrious collection. The man who goes through the streets of a town and puts down the names of each of the inhabitants, with their occupations and their street number, acquires material of which he is the author.

*Jeweler's Circular Publishing Co. v. Keystone Publishing Co.*, 281 F. 83, 88 (2d Cir.) (emphasis added), *cert. denied*, 259 U.S. 581 (1922).

The doctrine was affirmed in *Leon v. Pacific Tel. & Tel. Co.* where the court found infringement when the defendant copied information from the plaintiff's alphabetical telephone directory and produced a new directory arranged by telephone number. 91 F.2d 484, 487 (9th Cir. 1937). The court relied on the above quoted passage from *Jeweler's* in protecting the information. *Id.* at 486.

While these decisions preceded the 1976 revisions of the Copyright Act, subsequent decisions have also relied on the doctrine. See *Rockford Map Publishers, Inc. v. Directory Serv. Co.*, 768 F.2d 145, 149 (7th Cir. 1985) (after denying that protection can be based on effort, nevertheless protecting the information in a map by requiring a second compiler to assemble the information as if there had never been a first compilation and then allowing use of the first only as a check on error), *cert. denied*, 474 U.S. 1061 (1986); *Hutchinson Tel. Co. v. Fronteer Directory Co.*, 770 F.2d 128, 131-32 (8th Cir. 1985) (relying on "sweat of the brow" in finding a white pages directory copyrightable). See also *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 663 F. Supp. 214, 218 (D. Kan. 1987), *aff'd without opinion*, 916 F.2d 718 (10th Cir. 1990), *rev'd*, 111 S.Ct. 1282 (1991).

Even before *Feist*, "sweat of the brow" faced considerable criticism. See, e.g., *Shipley & Hay*, *supra* note 16, at 139-41 (1984); 1 NIMMER & NIMMER, *supra* note 14, § 3.04, 3-22 to 3-23. *But see* Ginsburg, *supra* note 10, at 1875-81 (analyzing and defending the historical basis for protection based on the author's labor).

27. See *Eckes v. Card Prices Update*, 736 F.2d 859, 862-63 (2d Cir. 1984) (noting that "subjective selection and arrangement of information" merits protection); *Miller v. Universal City Studios, Inc.*, 650 F.2d 1365, 1369 (5th Cir. 1981) (arguing that a compilation copyright is "properly viewed as resting on the originality of the selection and arrangement of the factual material, rather than on the industriousness of the efforts to develop the information"). See also *Worth v. Selchow & Righter Co.*, 827 F.2d 569, 573 (9th Cir. 1987) (Ninth Circuit rejection of *Leon*), *cert. denied*, 485 U.S. 977 (1988); *Financial Info., Inc. v. Moody's Investors Serv., Inc.*, 808 F.2d 204, 207 (2d Cir. 1986), *cert. denied*, 484 U.S. 820 (1987); *Cooling Sys. & Flexibles, Inc. v. Stuart Radiator, Inc.*, 777 F.2d 485, 491 (9th Cir. 1985); *Southern Bell Tel. & Tel. Co. v. Associated Tel. Dir. Publishers*, 756 F.2d 801, 809 (11th Cir. 1985).

northwest Kansas, annually published a standard telephone directory consisting of both white and yellow pages.<sup>28</sup> Feist Publications, Inc. specialized in area-wide telephone directories covering a larger geographic area than those provided by most telephone companies.<sup>29</sup> Because it is not a telephone company and had no expedient means of obtaining subscriber information, Feist attempted to purchase licenses from each of the eleven telephone companies operating in various parts of northwest Kansas for use of their white pages listings. Although only Rural refused,<sup>30</sup> omitting its listings would have left a significant gap in Feist's directory; therefore, Feist ignored Rural's refusal and incorporated its listings. After removing entries from locales falling outside the scope of its area-wide directory, Feist hired personnel to verify and expand upon<sup>31</sup> the approximately 5,000 entries that remained. At the end of this process a substantial number of listings in Feist's directory were still identical to those in Rural's.<sup>32</sup>

Rural sued for copyright infringement,<sup>33</sup> arguing that Feist could not use Rural's copyrighted white pages in creating its own directory. Feist responded that telephone directories were not subject matter worthy of a copyright.<sup>34</sup> The district court granted summary judgment to Rural based on numerous lower court decisions which had consistently held telephone directories to be copyrightable.<sup>35</sup> In an unpublished opinion, the Court of Appeals for the Tenth Circuit affirmed for essentially the same reasons as the district court.<sup>36</sup> On certiorari, the Supreme Court unanimously reversed.<sup>37</sup>

---

28. See *Feist*, 111 S.Ct. at 1286.

29. Using these directories reduces the need to consult multiple directories or call directory assistance. The disputed Feist directory covered 11 different telephone service areas in 15 counties and contained 46,878 white pages listings—compared to Rural's approximately 7,700 listings. See *id.*

30. See *id.*

31. As a result, while most of Rural's listings included just a subscriber's name, town, and telephone number, Feist's typical listing also included the street address. See *id.*

32. "Notwithstanding these additions, however, 1,309 of the 46,878 listings in Feist's 1983 directory were identical to listings in Rural's 1982-1983 white pages. Four of these were fictitious listings that Rural had inserted into its directory to detect copying." *Id.* at 1286-87 (citations omitted).

33. *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 663 F.Supp. 214 (D. Kan. 1987), *aff'd without opinion*, 916 F.2d 718 (10th Cir. 1990), *rev'd*, 111 S.Ct. 1282 (1991).

34. See *id.* at 217.

35. See *id.* at 218.

36. See appendix to Pet. for Cert. at 4a, *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282 (1991) (No. 89-1909), judgment order reported at 916 F.2d 718 (10th Cir. 1990).

37. See *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282 (1991). Seven other justices joined in the opinion written by Justice O'Connor. Justice Blackmun did not join in the opinion but concurred in the judgment.

### B. Elimination of "Sweat of the Brow"

The conflict in *Feist* centered around the tension between two traditional propositions of copyright law.<sup>38</sup> First, facts are not copyrightable. "That there can be no valid copyright in facts is universally understood. The most fundamental axiom of copyright law is that '[n]o author may copyright his ideas or the facts he narrates.'"<sup>39</sup> Meanwhile, compilations of facts generally are copyrightable.<sup>40</sup> Even compilations consisting exclusively of facts are potentially within the scope of copyright.<sup>41</sup>

Because "sweat of the brow" provided copyright protection to pure facts,<sup>42</sup> the Court decided to eliminate it as an acceptable theory of protection. As a first step in discrediting "sweat of the brow," the Court established that originality is the *sine qua non* of copyright:

To qualify for copyright protection, a work must be original to the author. Original, as the term is used in copyright, means only that the work was independently created by the author (as opposed to copied from other works), and that it possesses at least some minimal degree of creativity. To be sure, the requisite level of creativity is extremely low; even a slight amount will suffice.<sup>43</sup>

A constitutional basis for this originality requirement emanated from two early Supreme Court opinions.<sup>44</sup> The Court found additional support for

38. *See id.* at 1287.

39. *Id.* (quoting *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 556 (1985)). *But see* Ginsburg, *supra* note 10, at 1913-16 (arguing that the fact-expression distinction does not merit the same role in copyright law as the related idea-expression distinction).

40. *See* 111 S.Ct. at 1287.

41. *See id.* *See also* 1 NIMMER & NIMMER, *supra* note 14, § 2.04[B], 2-41 to 2-42.

42. *See supra* note 26.

43. 111 S.Ct. at 1287 (citations omitted).

44. *See id.* at 1288. The Constitution authorizes Congress to "secur[e] for limited Times to Authors . . . the exclusive Right to their respective Writings." U.S. CONST. art. I, § 8, cl. 8. "[W]hile the word *writings* may be liberally construed . . . it is only such as are *original*, and are founded in the creative powers of the mind." The Trade-Mark Cases, 100 U.S. 82, 94 (1879). The *Feist* Court construed this as explaining that "originality requires independent creation plus a modicum of creativity." 111 S.Ct. at 1288. Another opinion defined "author" to mean "he to whom anything owes its origin; originator; maker." *Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53, 58 (1884) (internal quotations omitted). The *Feist* Court, 111 S.Ct. at 1288, found emphasis of the "creative component of originality" in the *Burrow-Giles* Court's limitation of copyright to "original intellectual conceptions of the author," 111 U.S. at 58, and the burden it placed on the author to prove "the existence of those facts of originality, of intellectual production, of thought, and conception." *Id.* at 59-60.

this proposition in recent cases<sup>45</sup> and scholarly works.<sup>46</sup>

"Sweat of the brow" has no foundation under the Copyright Act of 1976 according to *Feist* because it imposes no originality requirement.<sup>47</sup> By basing protection solely on the author's efforts, the "sweat of the brow" courts offered protection for the facts themselves. However, as the Court noted, "[n]o one may claim originality as to facts."<sup>48</sup> Therefore, the originality requirement precludes protection under "sweat of the brow." The doctrine actually arose from some lower courts' misinterpretation of ambiguities in the Copyright Act of 1909.<sup>49</sup> Congress attempted to remedy this situation in the Copyright Act of 1976 in a number of ways. First, the Act explicitly limited copyright protection only to "original works of authorship."<sup>50</sup> Next, it added section 102(b) which prohibits protection of "any idea, procedure, process, system, method of operation, concept, principle, or discovery"<sup>51</sup> and which has

45. See *Goldstein v. California*, 412 U.S. 546, 561-62 (referring with approval to the definitions in *Burrow-Giles* and *The Trade-Mark Cases*), *reh'g denied*, 414 U.S. 883 (1973); *Miller v. Universal City Studios, Inc.*, 650 F.2d 1365, 1368 (5th Cir. 1981) (noting that originality is "the premise of copyright law").

46. See L. Ray Patterson & Craig Joyce, *Monopolizing the Law: The Scope of Copyright Protection for Law Reports and Statutory Compilations*, 36 UCLA L. REV. 719, 763 n.155 (1989) (noting that "[t]he originality requirement is constitutionally mandated for all works"). See also 1 NIMMER & NIMMER, *supra* note 14, § 1.06[A] ("originality is a statutory as well as a constitutional requirement"); *id.* § 1.08[C][1] ("a modicum of intellectual labor . . . clearly constitutes an essential constitutional element"); Patterson & Joyce, *supra*, at 759-60 & n.140. *But see* Ginsburg, *supra* note 10, at 1873-74 (arguing that this creativity requirement is a product of the "mid to late nineteenth century" and that prior decisions have based protection on labor and thus the creativity requirement is not constitutionally mandated).

47. See 111 S.Ct. at 1295. *But see* Ginsburg, *supra* note 10, at 1895-96 (arguing that "[i]t is not clear that the statute demands this rejection of the 'sweat' test of authorship").

48. 111 S.Ct. at 1288 (quoting 1 NIMMER & NIMMER, *supra* note 14, § 2.11[A], 2-157). The Court noted that the "distinction is one between creation and discovery: the first person to find and report a particular fact has not created the fact; he or she has merely discovered its existence." The Court gives the example of census-takers who copy population figures from the world around them rather than create the data themselves. *But see* Beck, *supra* note 16, at 8-9 n.28 (criticizing this view from an epistemological standpoint).

49. 111 S.Ct. at 1290-91. While, in the Court's view, §§ 3-4 of the 1909 Act incorporated the originality requirement, they did so implicitly and were ambiguous. *Id.* The "sweat of the brow" courts ignored §§ 3-4 and focused their attention on § 5, a technical provision listing 14 types of copyrightable works. *Id.* at 1291. Because one of these categories included compilations, the "sweat of the brow" courts mistakenly inferred that compilations were copyrightable per se, "without any further or precise showing of original—personal—authorship." *Id.* (quoting Ginsburg, *supra* note 10, at 1895).

50. 17 U.S.C. § 102(a) (1988).

51. *Id.* § 102(b). This section is the statutory basis for the idea-expression dichotomy discussed *infra* Section IV.

been "universally understood" to prevent copyright in facts.<sup>52</sup> Finally, the tripartite structure<sup>53</sup> of the Act's definition of "compilation"<sup>54</sup> indicates an originality requirement and thereby eliminates any notion of *per se* copyrightability.<sup>55</sup>

### C. Selection and Arrangement

In addition to eliminating "sweat of the brow," the Court restated the accepted basis for protecting factual compilations. Factual compilations, even those not containing any protectable written expression,<sup>56</sup> may merit protection because, according to *Feist*, the compilation author typically makes sufficiently original choices as to selection and arrangement of the data.<sup>57</sup> To qualify, these choices need only be made independently by the compiler, as opposed to copied from other sources, and include a minimal level of creativity;<sup>58</sup> novelty is not required.<sup>59</sup> The first requirement is inherent in the concept of originality—something is not original if it has been copied from another source. As for the second requirement, the Court noted that "the requisite level of creativity is extremely low [and] . . . [t]he vast majority of works make the grade quite easily, as they possess some creative spark, 'no matter how crude, humble or obvious' it might be."<sup>60</sup> The facts need not be presented in an unusual or innovative way,<sup>61</sup> however, the selection and arrangement "cannot be so mechanical or routine as to require no creativity whatso-

---

52. 111 S.Ct. at 1293. See *Harper & Row, Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 556 (1985).

53.

The statute identifies three distinct elements and requires each to be met for a work to qualify as a copyrightable compilation: (1) the collection and assembly of pre-existing material, facts, or data; (2) the selection, coordination, or arrangement of those materials; and (3) the creation, by virtue of the particular selection, coordination, or arrangement, of an 'original' work of authorship.

111 S.Ct. at 1293.

54. 17 U.S.C. § 101 (1988). See *supra* note 15.

55. See *supra* note 49.

56. See 111 S.Ct. at 1289. See also *Harper & Row*, 471 U.S. at 547.

57. See 111 S.Ct. at 1289.

58. See *id.* See also 1 NIMMER & NIMMER, *supra* note 14, §§ 2.11[D], 3.03; Denicola, *supra* note 17, at 523 n.38.

59. See 111 S.Ct. at 1294.

60. *Id.* at 1287 (quoting 1 NIMMER & NIMMER, *supra* note 14, § 1.08[C][1]).

61. See *id.* at 1296.

ever."<sup>62</sup> While the creativity standard is low, some works will be unable to meet it.<sup>63</sup>

The sole issue in the Court's analysis of the particular facts of *Feist* was determining whether Rural's selection and arrangement was original.<sup>64</sup> To establish infringement, Rural needed to show that it owned a valid copyright and that Feist had copied protected elements of its compilation.<sup>65</sup> Because Feist had conceded that Rural held a valid copyright, Rural met the first element.<sup>66</sup> As to the second element, Rural's directory consisted of two potentially protectable elements: subscriber information and Rural's selection and arrangement of that information. While Rural was the first to discover and report the subscriber information in its directory, this data consisted of uncopyrightable facts and was not separately protectable.<sup>67</sup> The only potentially protectable element that Feist could have copied was Rural's selection and arrangement.

The Supreme Court held that Rural's "garden-variety" directory was "entirely typical[,] . . . devoid of even the slightest trace of creativity," and therefore lacking sufficient originality to qualify for copyright protection.<sup>68</sup> The mere publication of basic information about each person applying for telephone service was not a sufficiently original selection of data.<sup>69</sup> Also, its arrangement of this information in alphabetical order lacked any semblance of creativity.<sup>70</sup> Rural's directory was the essence of unoriginality: "Given that some works must fail, we cannot imagine a more likely candidate."<sup>71</sup>

---

62. *Id.* See Patterson & Joyce, *supra* note 46, at 760 n.144 (noting that the creativity requirement is "not without effect").

63. See 111 S.Ct. at 1294. See also *Bleistein v. Donaldson Lithographing Co.*, 188 U.S. 239, 251 (1903).

64. See 111 S.Ct. at 1296.

65. See *id.*

66. See *id.*

67. See *id.* See also 1 NIMMER & NIMMER, *supra* note 14, § 2.03[E].

68. 111 S.Ct. at 1296.

69. See *id.* at 1297. "This is 'selection' of a sort, but it lacks the modicum of creativity necessary to transform mere selection into copyrightable expression." *Id.*

70. See *id.* "[A]rranging names alphabetically in a white pages directory . . . is an age-old practice, firmly rooted in tradition and so commonplace that it has come to be expected as a matter of course." *Id.*

71. *Id.*

### III. PROTECTION AFTER *FEIST*

While *Feist* establishes that originality in selection and arrangement precludes copying and that "mechanical" selection and arrangement is insufficient to meet the creativity standard, the Court's analysis leaves two critical questions unanswered. First, what constitutes creativity in a selection or arrangement? Second, if a selection or arrangement is original, what sort of copying does it prevent? Because *Feist* provides little guidance as to these issues, an examination of other opinions is informative.

#### A. *When is a Selection or Arrangement Creative?*

##### 1. *The Current Standard*

Two recent opinions in the Second and Eleventh Circuits have required the compiler to make subjective decisions in order to pass the *Feist* creativity standard.<sup>72</sup> Creative selection requires either the conscious exclusion from the compilation of some data that meets the objective criteria of the compilation or subjective decisions in determining these criteria. Creative arrangement requires the compiler to arrange data in some manner other than in a mechanical or objective fashion. This notion is consistent with other "selection and arrangement" opinions.<sup>73</sup>

The Second Circuit examined the *Feist* creativity standard in determining whether infringement of a classified business directory for use by New York City's Chinese-American community had occurred.<sup>74</sup> At first, the court seemed to disregard the creativity standard when it noted that originality in practice is little more than a prohibition on copying by

---

72. See *Key Publications, Inc., v. Chinatown Today Publishing Enters., Inc.*, 1991 U.S. App. LEXIS 22250 (2d Cir. Sept. 23, 1991); *Bellsouth Advertising & Publishing Corp. v. Donnelley Info. Publishing, Inc.*, 933 F.2d 952 (11th Cir. 1991).

73. See, e.g., *Eckes v. Card Prices Update*, 736 F.2d 859, 862-63 (2d Cir. 1984) (finding the exercise of judgment in selecting a set of "premium" baseball cards worthy of copyright protection). See also 1 NIMMER & NIMMER, *supra* note 14, § 2.04[B], n.15.1 (examining opinions which found originality in selection and emphasizing the necessity of "subjective judgment"); *Ginsburg*, *supra* note 10, at 1896 (describing selection and arrangement as "a test of subjective authorship"); *Hicks*, *supra* note 12, at 1005 (noting that "under the subjective selection standard expression is found in the actions taken to collect information") (emphasis added).

74. See *Key Publications*, 1991 U.S. App. LEXIS 22250.

the initial compiler.<sup>75</sup> However, the court then noted that original selection involves judgment in choosing which facts from a body of data to include in a compilation.<sup>76</sup> The plaintiff's exclusion of businesses that it did not think would remain open very long was held to be sufficiently original selection.<sup>77</sup> Original "[a]rrangement 'refers to the ordering or grouping of data into lists or categories that go beyond the mere mechanical grouping of data as such, for example, the alphabetical, chronological, or sequential listings of data.'"<sup>78</sup> While the plaintiff's directory was in a format similar to most other classified directories,<sup>79</sup> the "arrangement [was] in no sense mechanical, but involved creativity on the part of [the plaintiff] in deciding which categories to include and under what name."<sup>80</sup> Thus, the plaintiff's selection and naming of categories was sufficiently creative to constitute original arrangement and involved the minimal thought necessary to meet the originality requirement.<sup>81</sup>

The Eleventh Circuit used similar criteria in finding original selection and arrangement in a yellow pages directory.<sup>82</sup> Based on the plaintiff's selection and demarcation of geographic boundaries for each of its directories, it listed a particular business in one or more of these compilations.<sup>83</sup> Additional acts of selection noted by the court included selecting a date on which a directory could no longer be modified and creating classifications for the various businesses.<sup>84</sup> The plaintiff's directories were arranged according to both these classifications and the geographical scope selected for each directory.<sup>85</sup> *Bellsouth*, then, also required subjectivity as a prerequisite for finding creativity. While the geographic aspect of selection and arrangement borders on the mechanical, the court relied on the subjective arrangement of businesses into classifications in its finding of originality.

Though reasonable on its face, a closer examination of this subjective

---

75. See *id.* at \*8-9. See also Denicola, *supra* note 17, at 520-21; 1 NIMMER & NIMMER, *supra* note 14, § 2.01[B].

76. See 1991 U.S. App. LEXIS 22250 at \*9. See also *Eckes*, 736 F.2d at 862-63 (finding selectivity in the designation of 5,000 cards as "premium" from among 18,000 baseball cards); Patry, *supra* note 17, at 57.

77. See 1991 U.S. App. LEXIS 22250 at \*10.

78. *Id.* at \*12 (quoting Copyright Office, Guidelines for Registration of Fact-Based Compilations 1 (Rev. Oct. 11, 1989), *quoted in* Patry, *supra* note 17, at 60).

79. See *id.*

80. *Id.* at \*13.

81. See *id.*

82. *Bellsouth Advertising & Publishing Corp. v. Donnelley Info. Publishing, Inc.*, 933 F.2d 952 (11th Cir. 1991).

83. See *id.* at 957.

84. See *id.*

85. See *id.* at 958.

creativity standard reveals an underlying flaw. The courts purport to find subjectivity in a compiler's decisions about selection and arrangement; however, these decisions are based on functional considerations and are not truly subjective.<sup>86</sup> For example, the selection of categories in *Key Publishing*, though subject to the compiler's judgment, was not based on any notion that the categories chosen were aesthetically pleasing. Rather, the selection of categories was based on the compiler's notion that they would aid the directory's users in locating information. The exclusion of certain businesses from the directory was likewise based on a desire to avoid including questionable listings, thereby increasing the directory's functional value rather than asserting anyone's individuality. Thus, the opinions finding creative expression in a compiler's selection and arrangement appear to find expression where none actually exists.<sup>87</sup>

## 2. Application of the Standard to Databases

The *Feist* standard is particularly troublesome for databases, as opposed to other factual compilations, for three reasons. First, most modern databases are comprehensive, and this impedes their protection under the subjective selection standard just discussed. Second, the traditional notion of "arrangement" is difficult to apply to computer databases. Third, the automated nature of databases enhances their functionality and places further strain on the dubious notion that selection and arrangement are subjective.

The subjectivity requirement for original selection frustrates database development because developers strive to create comprehensive rather than subjectively limited databases.<sup>88</sup> Two primary advantages offered by computers are their storage capacity and their ability to quickly access particular data from within a complex database.<sup>89</sup> To effectively utilize these characteristics, a database developer generally tries to make an extremely comprehensive database. Rather than subjectively leaving

---

86. Subjective is defined as "[p]roceeding from or taking place within an individual's mind and unaffected by the outside world." THE AMERICAN HERITAGE DICTIONARY 677 (Office ed. 1983). Because these decisions are based on functional considerations, they are certainly affected by the outside world and are not subjective.

87. See Ginsburg, *supra* note 10, at 1868 & n.13 (noting that the "[p]ersonality-based characterizations of many [factual compilations] seem contrived").

88. See IIA/ADAPSO Brief, *supra* note 3, at 19-20 n.16 (arguing that under the original selection standard "copyright protection would be afforded to compilations in inverse proportion to their comprehensiveness, which is to say (in many cases) to their commercial value and usefulness"). See also *Financial Info., Inc. v. Moody's Investors Serv., Inc.*, 751 F.2d 501, 507 (2d Cir. 1984) (basing selectivity on the omission of facts).

89. See Hicks, *supra* note 12, at 995.

information out of a database, a developer strives to include as much information as possible.<sup>90</sup> Indeed, a more comprehensive database is more useful and desirable to most users. A database's scope is also generally based on objective criteria<sup>91</sup> so that its users can make their own subjective decisions about which data interests them.<sup>92</sup> Under the present standard, a comprehensive database constructed on objective criteria is not protected.<sup>93</sup> This leads to the troubling result that a more useful database is less likely to be protected than its less useful counterpart.<sup>94</sup>

Another weakness in current protection schemes is created by the fact that flexible arrangement, while making a database more useful, reduces the likelihood of its protection. Initially, the concept of "arrangement" as embodied in a yellow pages directory, for example, is difficult to envision in the context of a database.<sup>95</sup> This difficulty has led one court to note that a database stores information "without arrangement and form";<sup>96</sup> thus, protection of databases based on original arrangement is unlikely. This lack of arrangement is, however, a very positive feature

---

90. See *id.* at 1006 (noting that most compilations "stand out for their exhaustiveness and usually contain components selected on the basis of objective, not subjective, criteria").

91. See *id.*

92. For example, rather than providing only the name and address of each person living in Boston who is over 50 and earns more than 50,000 dollars per year, a database would provide the name, address, age, and income of each person living in Boston. A user could then examine data meeting any particular age or income criteria that she desires.

93. The selection used in the database discussed *supra* note 92 would not be protected. See Hicks, *supra* note 12, at 1006 (noting that a compilation listing "all known inhabitants of a city . . . would exhibit insufficient subjectivity"). But see *id.* at 1006-07 (arguing that subjectivity is involved even in the selection of objective criteria).

Of course, the developer of that database could argue that its choice of fields, i.e. name, address, age, and income, constitutes original selection, but this selection would likely be deemed "entirely typical" and unworthy of protection. See *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282, 1296 (1991) (finding that the selection of name, town, and telephone number "could not be more obvious").

94. The database discussed *supra* note 92 would not be protected while a similar database containing information about only those people over 50 years old who earn more than 50,000 dollars per year would be marginally more likely to be protected.

95. The physical arrangement of data within a database is important to the efficiency of the database system. See Hicks, *supra* note 12, at 1014. However, this concept of arrangement is purely functional, and protection of any original method of physically arranging data would surely be barred by the idea-expression dichotomy as embodied in § 102(b). See *supra* note 51 and accompanying text; *infra* Section IV.

96. *National Business Lists, Inc. v. Dun & Bradstreet, Inc.*, 552 F. Supp. 89, 97 (N.D. Ill. 1982). See also Denicola, *supra* note 17, at 531 (arguing that "there is no particular arrangement to protect"). But see Hicks, *supra* note 12, at 1022 (criticizing this view as "factually incorrect").

of databases since it allows users to arrange data in the format they find most useful.<sup>97</sup> Next, even if a developer can somehow establish the presence of an original arrangement, this prong of the *Feist* standard offers virtually no protection to databases. A subsequent developer can easily rearrange the data and thus avoid infringement without impeding the utility of the subsequent database.<sup>98</sup>

The enhanced functionality of databases places additional strain on the notion that there is subjectivity in their selection and arrangement. Though the selection and arrangement of traditional compilations are generally based on functional considerations,<sup>99</sup> the compiler can at least argue that the arrangement attempts to be visually pleasing or that subjective choices were made in the selection of data. A database developer, however, can make no such claims. Authors of more traditional compilations are generally limited in the amount of information that they can include in their works and thus must make arguably subjective choices. However, because the software portion of a database system allows easy manipulation of the database, a developer does not face the dilemma of providing too much information and thus data meeting the objective criteria of the database need not be excluded. Therefore, database developers generally will not have made these supposedly subjective decisions which more traditional authors often reach.

## B. What Sort of Copying is Prevented?

### 1. The Current Rule

Because the copyright in a factual compilation is "thin,"<sup>100</sup> the protection offered to an original selection or arrangement is limited to the components that originated with the author.<sup>101</sup> A selection or arrangement can be original and protectable, but facts never become original through

---

97. See Hicks, *supra* note 12, at 1014 ("The database compiler leaves the determination of the manner in which data is viewed to the discretion of the user within the confines of the software.")

98. See *id.* at 1023 (noting that rearrangement will avoid infringement and that the ease with which a technically proficient party can rearrange the data "leads to shallow protection").

99. See *supra* notes 86-87 and accompanying text.

100. *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282, 1289 (1991).

101. See *id.* See also Ginsburg, *supra* note 10, at 1868 & n.12; Patterson & Joyce, *supra* note 46, at 800-02.

association.<sup>102</sup> Therefore, an original selection or arrangement offers minimal protection:

[A] subsequent compiler remains free to use the facts contained in another's publication to aid in preparing a competing work, so long as the competing work does not feature the same selection and arrangement. . . . "[T]he very same facts and ideas may be divorced from the context imposed by the author, and restated or reshuffled by second comers, even if the author was the first to discover the facts or to propose the ideas."<sup>103</sup>

Copyright protects only the selection and arrangement of a factual compilation—"[i]n no event may copyright extend to the facts themselves."<sup>104</sup>

Because the Court in *Feist* found no originality in Rural's selection and arrangement, it did not reach the issue of what amount of copying will actually infringe a compilation copyright. An original selection or arrangement will certainly prevent verbatim copying of an entire compilation, but it will not prevent copying of an isolated fact. At what point between these extremes does copying become impermissible?

The Second Circuit's decision in *Key Publications*<sup>105</sup> clarified the extent of copying prevented by an original selection and arrangement. The court noted that while *Feist* could be interpreted as limiting compilation protection to only the prevention of verbatim copying, it had refused to "read *Feist* in such a broad and self-defeating fashion."<sup>106</sup> Subsequent compilers cannot simply modify a copyrighted compilation with a single fact and thereby avoid infringement suits.<sup>107</sup> "Such a result would

---

102. See 111 S.Ct. at 1289. See also Patry, *supra* note 17, at 64; Patterson & Joyce, *supra* note 46, at 776.

103. 111 S.Ct. at 1289 (quoting Ginsburg, *supra* note 10, at 1868). Thus, "each subsequent compiler [need not] start from scratch and is [not] precluded from relying on research undertaken by another. Rather, the facts contained in existing works may be freely copied because copyright protects only the elements that owe their origin to the compiler . . ." *Id.* at 1295 (citations omitted).

104. *Id.* at 1290.

105. *Key Publications, Inc., v. Chinatown Today Publishing Enters., Inc.*, 1991 U.S. App. LEXIS 22250 (2d Cir. Sept. 23, 1991). See *supra* notes 73-81 and accompanying text.

106. 1991 U.S. App. LEXIS 22250 at \*14. See *Kregos v. Associated Press*, 937 F.2d 700, 710 (2d Cir. 1991) (noting that the defendant's selection of data must vary from the plaintiff's by more than a trivial degree to avoid infringement).

107. See 1991 U.S. App. LEXIS 22250 at \*14-15. See also *Harper & Row Publishers, Inc. v. Nation Enters.*, 471 U.S. 539, 583 n.5 (1985) (Brennan, J., dissenting).

render the copyright of a compilation meaningless."<sup>108</sup> Although the copyright in a factual compilation is "thin," it is not "anorexic."<sup>109</sup> The Second Circuit showed, then, that copyright protection in a factual compilation goes beyond the prevention of verbatim copying.

A plaintiff must prove "substantial similarity between those elements, and only those elements, that provide copyrightability to the allegedly infringed compilation" to establish infringement.<sup>110</sup> Because protection is limited to the original selection and arrangement provided by the compiler, similarity as to those elements must be proven. Thus, the plaintiff in *Key Publishing* had to show substantial similarity between the two directories with regard to their arrangements of categories or selections of businesses.<sup>111</sup>

Noting that "[t]he key issue is not whether there is overlap or copying but whether the organizing principle guiding the selection of businesses for the two publications is in fact substantially similar,"<sup>112</sup> the court found no substantial similarity in the selections of data. Fifteen hundred of the 2,000 listings in the defendants' directory duplicated listings in the plaintiff's directory, and these duplications resulted from intentional copying by the defendants.<sup>113</sup> However, the defendants had copied just seventeen percent (1,500/9,000) of the listings in the plaintiff's directory and had not duplicated in total any category containing more than a few

---

108. 1991 U.S. App. LEXIS 22250 at \*14.

109. *Id.* at \*15.

110. *Id.* Proof of "substantial similarity" is a basic requirement in establishing infringement. *Id.* See, e.g., *Business Trends Analysts, Inc. v. Freedonia Group, Inc.*, 887 F.2d 399, 402 (2d Cir. 1989); *Eckes v. Card Prices Update*, 736 F.2d 859, 863 (2d Cir. 1984). See also 3 NIMMER & NIMMER, *supra* note 14, § 13.03[A] ("Just as copying is an essential element of infringement, so substantial similarity . . . is an essential element of copying.") (footnotes omitted).

111. See 1991 U.S. App. LEXIS 22250 at \*16; *supra* notes 72-79 and accompanying text. "[T]he components of a compilation are generally in the public domain, and a finding of substantial similarity or even absolute identity as to matters in the public domain will not suffice to prove infringement." 1991 U.S. App. LEXIS 22250 at \*15. See *Bellsouth Advertising & Publishing Corp. v. Donnelley Info. Publishing, Inc.*, 933 F.2d 952, 957 (11th Cir. 1991) ("[T]here must be a substantial appropriation of the original format of the compilation to constitute the copying of protected material."); *Kregos v. Associated Press*, 937 F.2d 700, 709 (2d Cir. 1991) (compilation author is only protected against infringement of the protectable features of the compilation). See also 3 NIMMER & NIMMER, *supra* note 14, § 13.03[B][2][b].

112. 1991 U.S. App. LEXIS 22250 at \*19. Thus, substantial similarity is limited to the organizing principle of selection. However, protection of this principle is limited by the idea-expression dichotomy, which limits protection to the expression of an idea and denies protection to the idea itself. See *infra* Section IV.

113. See 1991 U.S. App. LEXIS 22250 at \*18.

listings.<sup>114</sup> The defendants' selection in *Key Publishing* did not infringe because there will always be significant overlap between classified directories for a given community.<sup>115</sup> Infringement would have been found, though, if the defendants had "exactly duplicated a substantial designated portion" of the plaintiff's directory.<sup>116</sup> Thus, an original selection will allow significant copying from the original work so long as the copy does not duplicate a substantial portion of the original.

The court's substantial similarity test for the "arrangement" prong of the *Feist* standard is also troubling. The court in *Key Publishing* ruled that, "as a matter of law," the arrangements in the two compilations were not substantially similar.<sup>117</sup> Only three of the twenty-eight categories in the defendants' directory duplicated any of the 260 categories in the plaintiff's directory.<sup>118</sup> Though many of the listings appearing in both directories appeared under similar headings, the court distinguished the arrangement of categories from the placement of listings within categories and noted that the latter is "the sort of mechanical task that does not merit copyright protection."<sup>119</sup> This statement by the court is troubling because the placement of listings within categories is no less subjective than any other aspect of a classified directory's arrangement. The categories in a classified directory are arranged in alphabetical order, clearly a mechanical process.<sup>120</sup> Likewise, the name chosen for a particular category is generally one that readily comes to mind when one envisions the product or service offered by the businesses listed within that category.<sup>121</sup> Any effort by a compiler to select creative category names would substantially diminish the utility of the directory.<sup>122</sup> Thus, while the placement of listings within a category is not particularly subjective, neither are the other decisions about arrangement in a directory.

Meanwhile, the Eleventh Circuit has found infringement of a yellow

---

114. *See id.* at \*18-19.

115. *See id.* at \*19.

116. *Id.* at \*23. *See Eekes v. Card Prices Update*, 736 F.2d 859, 863 (2d Cir. 1984) (finding infringement when a directory duplicated the listing of 5,000 cards designated as "premium" in a copyrighted 18,000 card directory).

117. 1991 U.S. App. LEXIS 22250 at \*18.

118. *See id.* at \*17.

119. *Id.* at \*17-18.

120. *See Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282, 1297 (1991). *See also supra* note 70 and text accompanying note 78.

121. For instance, information about medical services is listed under generic headings such as "doctor" or "physician."

122. Listings for legal services are placed under "lawyer" or "attorney" rather than "bar-rister," "counselor," or "solicitor." *See ROGET'S INTERNATIONAL THESAURUS* 767 (4th ed. 1977). While the latter are acceptable headings, users of directories generally would not think to look under these more exotic headings. Moreover, these headings are still relatively common and of doubtful creativity.

pages directory when the defendant "substantially appropriated" the "overall format" of the plaintiff's directory.<sup>123</sup> By keying into a computer database the factual information from the plaintiff's directory as well as codes corresponding to the plaintiff's business classifications, the defendant retained the ability to reproduce the plaintiff's entire arrangement.<sup>124</sup> Thus, an original arrangement provides protection against little more than exact duplication of that arrangement.

## 2. *The Current Rule's Implications for Databases*

The prohibition against verbatim copying is of limited value to database developers. Provided that a particular database is sufficiently original for protection,<sup>125</sup> this rule will discourage only lazy copying of the entire database by unsophisticated users.<sup>126</sup> Also, any duplication by individuals that does occur will be difficult to detect and prevent.<sup>127</sup> In reality, any protection offered by this prohibition against copying for private use will result from an individual's decision that copying a database without purchasing it is morally wrong rather than from fear of the legal implications of copyright infringement. Parties who would copy a database and then sell multiple copies of it are the real target of copyright law. These parties have both the technical expertise and the motivation to make the minimal changes necessary to avoid infringement.

Protection based on an "original arrangement" shares the same weakness as the prohibition on verbatim copying. Any party with the requisite technical expertise can easily avoid liability by simply rearranging the data. The copying party will probably have rearranged the data already as a means of avoiding infringement based on verbatim copying.

Though it offers more protection than "original arrangement," "original selection" is also of limited value. Initially, a database developer must establish that the chosen selection was sufficiently subjective to

---

123. *Bellsouth Advertising & Publishing Corp. v. Donnelley Info. Publishing, Inc.*, 933 F.2d 952, 960 (11th Cir. 1991).

124. *See id.* at 959.

125. *See supra* Section III.A.2.

126. Many users will not have the technical expertise necessary to make even minor modifications of a database and, thus, will be unable to make anything other than a verbatim copy. *See Hicks, supra* note 12, at 1023.

127. Database developers have no means of examining individual computer systems to determine if they contain unsanctioned copies of their databases. Even if they had the ability and desire to do so, the costs of bringing individual infringement actions would rarely justify the expected returns.

merit protection.<sup>128</sup> This selection then has to clear the additional hurdles imposed by the idea-expression dichotomy as well as the doctrine of merger.<sup>129</sup> Even then, a potential infringer is only prevented from copying a "substantial portion" of the developer's selection.<sup>130</sup> For insight into the weakness of protection offered to original selections, consider an extensive database containing detailed marketing information about all residents of Massachusetts. If someone copied all records from this database relating to people who live in Boston, this would probably be infringement because a substantial component of the database has been copied.<sup>131</sup> However, if the party were able to convince a court that this copying was based on a new "subjective" theory of selection (for example, people living in urban areas), he would conceivably be permitted to market that database. While this new database would appeal to a smaller class of users, it would be no more useful to those users than the original database.<sup>132</sup> Most users would buy this new database instead of the old one because the new developer has incurred no costs in gathering the information and can sell it at a substantially lower price.<sup>133</sup>

Protection based on original selection and arrangement is inadequate for databases. The contribution offered by a database is the informational value of its collection of a comprehensive set of data, and much of the costs of database development result from this collection process. Protection based on selection and arrangement, though, prevents copying of only the format of the data rather than the data itself and thus fails to protect the database's main contribution.<sup>134</sup>

---

128. See *supra* Section III.A.

129. See *infra* Section IV.

130. See *supra* text accompanying notes 112-16.

131. See *Key Publications, Inc., v. Chinatown Today Publishing Enters., Inc.*, 1991 U.S. App. LEXIS 22250 at \*23 (2d Cir. Sept. 23, 1991).

132. Any use possible with the new database can be carried out with the old one by simply adding the additional parameter that the people live in Boston.

133. Another disturbing example occurs when a subsequent developer combines the databases of two or more previous developers. For example, imagine that developers A and B have developed sufficiently original databases concerning, respectively, hotels in the Eastern and Western United States. Compiler C could combine these databases and market a database covering the entire United States. While this combined database is marginally more useful, the new developer will be able to sell his database for significantly less than the price of either of the original databases and thereby eliminates any market for the original databases.

134. See Ginsburg, *supra* note 10, at 1869 (noting that arrangement in factual compilations "may bear little, if any, connection to the work's central importance as a source of information").

#### IV. THE IDEA-EXPRESSION DICHOTOMY AND THE UTILITARIAN NATURE OF DATABASES

While not discussed in this context by the *Feist* Court,<sup>135</sup> the idea-expression dichotomy and the copyright system's resultant hesitancy to protect utilitarian works further limit protection of databases. In the landmark case of *Baker v. Selden*,<sup>136</sup> the plaintiff sought protection for blank forms required in the system of bookkeeping explained in his book. The lower court had ruled in favor of the plaintiff. The Supreme Court reversed based on the notion that to do otherwise would grant patent-like protection of the bookkeeping system to the plaintiff without proof of novelty as required by patent law.<sup>137</sup> This notion evolved into the idea-expression dichotomy, now a traditional part of copyright law.

Copyright protects only the expression of a work and not its underlying idea.<sup>138</sup> A corollary doctrine, the doctrine of merger, provides that when the possible means of expression are limited, the idea "merges" into the expression and neither can be protected.<sup>139</sup> Therefore, if protection for a database prevents others from using the idea of that database, the doctrine will apply, and protection of the database will be prohibited.

The implications of this doctrine were highlighted in a recent Second Circuit decision which found that the idea of an "outcome predictive pitching form" for baseball had not merged with its expression.<sup>140</sup> Initially the court warned that the merger doctrine must be applied with caution: If it is applied too readily, it denies protection to protectable

---

135. The Court did discuss it in the context of prohibiting protection of facts. See *Feist Publications, Inc. v. Rural Tel. Serv. Co.*, 111 S.Ct. 1282, 1290 (1991). However, the Court never discussed the problem of protecting a factual compilation's underlying idea.

136. 101 U.S. 99 (1879).

137. 101 U.S. at 103. See 1 NIMMER & NIMMER, *supra* note 14, § 2.18[B], 2-198 to 2-199.

138. See 1 NIMMER & NIMMER, *supra* note 14, § 2.03[D]. This view of the doctrine results from a Supreme Court interpretation of *Baker v. Selden* as holding that copying an idea without copying its expression does not constitute infringement. See *Mazer v. Stein*, 347 U.S. 201, 217 (1954). See also 1 NIMMER & NIMMER, *supra* note 14, § 2.18[D]. The doctrine was eventually codified by the Copyright Act of 1976. See 17 U.S.C. § 102(b) (1988) (prohibiting protection of "any idea, procedure, process, system, method of operation, concept, principle, or discovery").

139. See *Kregos v. Associated Press*, 937 F.2d 700, 705 (2d Cir. 1991) (noting that "even expression is not protected in those instances where there is only one or so few ways of expressing an idea that protection of the expression would effectively accord protection to the idea itself"). See also *Educational Testing Servs. v. Katzman*, 793 F.2d 533, 539 (3d Cir. 1986); *Toro Co. v. R & R Prods. Co.*, 787 F.2d 1208, 1212 (8th Cir. 1986).

140. *Kregos*, 937 F.2d at 705-07.

forms of expression; if applied too sparingly, it protects ideas.<sup>141</sup> Next, the court noted that the idea embodied within a factual compilation or database must be carefully defined:

In one sense, every compilation of facts can be considered to represent a merger of an idea with its expression. Every compiler of facts has the idea that his particular selection of facts is useful. If the compiler's idea is identified at that low level of abstraction, then the idea would always merge into the compiler's expression of it. . . . Even with an idea formulated at a somewhat high level of abstraction, circumstances might occur where the realistic availability of differing expressions is so drastically limited that the idea can be said to have merged in its expression.<sup>142</sup>

Thus, application of the merger doctrine to factual compilations and databases "depends on the level of abstraction at which the idea is formulated."<sup>143</sup> The more generally an idea is defined, the less likely it is that the doctrine of merger will apply.<sup>144</sup>

This need for generality is at odds with the subjective selection requirement,<sup>145</sup> so that their interaction further inhibits protection of databases. To satisfy the subjective selection requirement, a database developer must provide specific instances where subjective decisions were made. This will require him to define the idea of the database in very specific terms.<sup>146</sup> Therefore, to meet the subjective selection

---

141. See *id.* at 705. Thus, "courts have been cautious in applying the merger doctrine to selections of factual information." *Id.* See *Educational Testing*, 793 F.2d at 540 (finding the doctrine inapplicable to the selection of test questions); *Toro Co.*, 787 F.2d at 1212 (finding the doctrine inapplicable to selection of data for numbering parts). However, it "has been applied on occasion to selections of categories of data." 937 F.2d at 705. See *Matthew Bender & Co. v. Kluwer Law Book Publishers, Inc.*, 672 F. Supp. 107, 112 (S.D.N.Y. 1987) (finding that the plaintiff's categories of data concerning personal injury awards were "the only sensible ones which could have been used to compile the data").

142. 937 F.2d at 706. While the court notes that the idea of factual compilations must be formulated at higher levels or they would never be copyrightable, this assumes that factual compilations should be subject to copyright protection. If a basic rule of copyright must be modified to protect factual compilations, perhaps this indicates that factual compilations should not be protected by copyright.

143. *Id.* at 707. For Judge Learned Hand's classic formulation of this "abstractions test," see *Nichols v. Universal Pictures Corp.*, 45 F.2d 119, 121 (2d Cir. 1930), *cert. denied*, 282 U.S. 902 (1931).

144. 937 F.2d at 707.

145. See *supra* text following note 85.

146. If subjectivity exists, it will be found in the details of the selection process. The general principle of selection is likely to be objective.

requirement, a developer must argue that the concept of his database is very specific, while at the same time arguing that it is general enough to avoid the doctrine of merger. As these issues arise only in the context of an infringement action, the developer must simultaneously make these arguments to the same court.

The *Kregos* court offered an alternative view of the doctrine of merger as preventing protection somewhere along a continuum between selections based on "matters of taste and personal opinion"<sup>147</sup> and those that are "the first step in an analysis that yields a precise result or even a better-than-average probability of some result."<sup>148</sup> Protection of the former selections imposes little risk that an idea will be protected while protection of the latter selections imposes significant risk that protection of ideas will occur.<sup>149</sup> *Kregos*' pitching form was "part way along" this continuum.<sup>150</sup> He went beyond stating that his selected characteristics were the most pertinent; he was implying that they were useful in predicting outcomes.<sup>151</sup> However, he had not provided a system for weighing those statistics, so his idea was sufficiently general to avoid the doctrine of merger.<sup>152</sup> Under this view, application of the doctrine depends on very difficult and arbitrary line-drawing by individual courts. This line-drawing is at the court's discretion; only in extreme cases will a court be obligated to rule a certain way.

The *Kregos* court's view of the doctrine offers little comfort to database developers. Initially, databases are designed to be comprehensive and objective so that users can make selections based on their own notions of taste and personal opinion.<sup>153</sup> Thus, the analysis tends to start with databases already near an extreme at which courts have denied protection. Next, like the hypothetical system in *Kregos* that weighed statistics, database systems generally produce useful results, and this is likely to make the databases themselves fall within the doctrine of merger. Finally, because of the arbitrariness of the courts' line-drawing, a developer can never be sure whether the doctrine will apply to eliminate protection for a particular database.

---

147. 937 F.2d at 707. See *Eckes v. Card Prices Update*, 736 F.2d 859 (2d Cir. 1984) (designation of certain baseball cards as "premium"); *Social Register Ass'n v. Murphy*, 128 F. 116 (C.C.D.R.I. 1904) (selection of prominent families for inclusion in a social directory).

148. 937 F.2d at 707.

149. See *id.*

150. *Id.*

151. See *id.*

152. See *id.* But see *id.* at 711-12 (Sweet, J., dissenting) (arguing that *Kregos*' idea should be defined more specifically and thus merger should apply).

153. See *supra* notes 90-91 and accompanying text.

## V. A NEW SYSTEM OF PROTECTION

Copyright offers extremely limited protection to databases. The only protection now available is accorded merely to the "selection and arrangement" of a database, rather than to its real contribution, the assembly of various pieces of information into a coherent and useful whole. Thus the underlying problem of copyright protection for databases is that no protection is provided for the results of data collection.

This resulting set of data should be protected. Copyright is supposed to promote the "progress of Science and Useful Arts."<sup>154</sup> Databases serve vital functions in modern society ranging from the aid of researchers in their quest for knowledge to helping businesses efficiently market their products,<sup>155</sup> and protection of databases surely falls within the purview of this constitutional mandate. Yet, by failing to protect the set of data comprising the main value of a database, copyright fails to protect the database. Without this protection, a database can be easily copied by competitors without risk of liability, so developers will either discontinue development<sup>156</sup> of worthwhile databases or substantially curtail access to these databases. In either event, useful sources of information will be removed from the public.

While it might seem that providing copyright protection to data would eliminate public access to vital information, a closer examination reveals that, in practical application, this notion has little basis. Initially, protection of the set of data must be distinguished from protection of individual facts. When developers seek protection of the data that they have collected, they are not concerned with legitimate uses of the data by their customers. Indeed, the very purpose of developing a database is to provide users with a worthwhile source of information. A developer's real concern is that a competitor will take the product of his efforts and use it to develop a competing product that, because of its reduced development costs, can be sold at a much lower price. Expansive protection of data would not prevent access to that data; it would simply prevent unjust enrichment by subsequent developers. In addition, concern that legitimate uses of data will be curtailed by expansive protection can be addressed through a doctrine of fair use.<sup>157</sup> Protection of data will

---

154. U.S. CONST. art. I, § 8, cl. 8. See *supra* note 14.

155. See *supra* note 4.

156. See Ginsburg, *supra* note 10, at 1907 (noting that "given the ease with which computers may copy and reorganize information, failure to protect the facts themselves deprives the compiler of a meaningful incentive to production") (footnote omitted).

157. The factors to be used in determining fair use under the current copyright statute are directly applicable to evaluating uses of data:

stimulate its dissemination because the knowledge that their data is protected will encourage developers to expand access to their databases.

Adequate protection of databases might conceivably be reached through modification of the present copyright system,<sup>158</sup> but this is not a feasible alternative. The present system adequately serves its primary function of protecting literary and artistic works. The wholesale modifications needed to provide sufficient protection to databases would probably create as many problems in protecting the standard subject matter of copyright as they would resolve for databases. History has proven that the chances of implementing any necessary modifications in a realistic time period are minimal.<sup>159</sup>

A better and more realistic alternative would be for Congress to provide a system of *sui generis* protection for databases and factual compilations.<sup>160</sup> This system could function as a federal misappropriation or unfair competition law. Thus, it could provide extensive protection against competing uses of information while allowing liberal use by the general public.

In many respects, this system would be similar to copyright. Initially, it would secure to the developer the same rights of copying, preparation of derivative works, and distribution as provided by copyright.<sup>161</sup> While copyright also provides for the rights of performance and display,<sup>162</sup> these would be unnecessary for protection of factual compilations. Like copyright, this system would limit protection through a fair use

- 
- (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;
  - (2) the nature of the copyrighted work;
  - (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
  - (4) the effect of the use upon the potential market for or value of the copyrighted work.

17 U.S.C. § 107 (1988). The first factor weighs against preventing innocent uses of data. Under the second factor, a more generous exception could be provided for factual compilations. The third factor could allow limited use by subsequent compilers and thus prevent unreasonable monopolization of the information. Finally, the fourth factor could be used to implement the real purpose of protecting factual compilations: prevention of copying by commercial competitors.

158. See Ginsburg, *supra* note 10, at 1916-18 (advocating modification of the copyright system to protect factual compilations); *id.* at 1924-36 (proposing protection subject to compulsory licensing).

159. For example, the Copyright Act of 1976 was the result of a process begun in 1955. See ALAN LATMAN ET AL., COPYRIGHT FOR THE NINETIES 10 (1989).

160. Precedent for this sort of protection already exists. See Semiconductor Chip Protection Act of 1984, 17 U.S.C. §§ 901-14 (1988) (providing *sui generis* protection to semiconductor chips).

161. See *id.* § 106 (1988).

162. See *id.*

provision. In fact, the fair use provision of the copyright statute would only need minor alterations to reflect the policy of allowing liberal use of information by the public while providing extensive protection against competing uses of the information.<sup>163</sup>

The system would also need to provide for a limited form of the idea-expression distinction.<sup>164</sup> This would prohibit protection of the idea underlying a database or factual compilation. Without this distinction, a compiler could prevent others from creating similar compilations even through reference to the original sources. However, this system of protection is based on a desire to protect a compiler's efforts in gathering information. Protection of the underlying idea goes beyond protection of this effort and thus exceeds the purpose of this system. To protect ideas would risk removing information from public access rather than just preventing subsequent compilers from stealing the fruits of the initial compiler's labors.

Ideally, the subject matter of this system should be limited to databases and factual compilations. Works like biographies and histories, though certainly worthy of some sort of protection, should not be protected by this proposal. Its purpose is to protect factual compilations because they now receive inadequate protection under copyright. Meanwhile, works like biographies and histories contain a large element of traditionally protectable expression and thus receive adequate protection under the copyright statute. Protection of the information contained in these works seems much more likely to impinge on the public's access to information and ideas.<sup>165</sup> Though this risk could probably be curtailed through a fair use analysis, to do so would simply complicate application of the system to its primary beneficiaries.

Noting that protection should be limited to factual compilations and databases is much easier than implementing this limitation. Rather than providing some sort of guideline that distinguishes between databases and works such as histories and biographies, a better alternative is to force the creator to choose between copyright protection or protection under this proposed system. While compilers and database developers would certainly select the expanded protection offered by this proposal, authors whose works include expression protectable by copyright would be less likely to do so. This proposal's fair use provision could be

---

163. See *supra* note 157.

164. See *supra* Section IV.

165. Works such as these are often written from varying perspectives and thus to foreclose use of the information contained in them risks silencing opposing viewpoints. Meanwhile, factual compilations provide information rather than some political interpretation of that information.

modified to further encourage authors of expressive works to seek copyright protection. Additional impetus could come from limitations on the term of protection.

The term of protection under this proposal should differ substantially from that of copyright. Copyright's provision for a term of life plus fifty years<sup>166</sup> is too long for works of this nature. Protection of information-based works, while promoting their creation, places a burden on their accessibility. So the term of protection must be long enough to encourage the development of works but not so long that access to information is unduly hindered. To this end, a term of ten years is suggested. This should provide enough time that a compiler will be able to realize adequate returns without jeopardizing public access. Finally, to account for the dynamic nature of information, developers should be encouraged to constantly update the information in their databases. This could be achieved by awarding a new term of protection to each version of a database.

---

166. See 17 U.S.C. § 302(a) (1988).

