Is There a Property Interest in Scientific Research Data

Richard H. Jones

Follow this and additional works at: https://scholarship.law.berkeley.edu/btlj

Recommended Citation

Link to publisher version (DOI)
https://doi.org/10.15779/Z38T67G

This Article is brought to you for free and open access by the Law Journals and Related Materials at Berkeley Law Scholarship Repository. It has been accepted for inclusion in Berkeley Technology Law Journal by an authorized administrator of Berkeley Law Scholarship Repository. For more information, please contact jcera@law.berkeley.edu.
IS THERE A PROPERTY INTEREST IN SCIENTIFIC RESEARCH DATA?

BY RICHARD H. JONES †

When scientists perform experiments or make observations, they record information in the form of computer printouts, handwritten lists of data, and photographs. This kind of research data is commonly thought to be copyrightable. For example, technical and professional journals routinely attach copyright notices to articles reporting data. The provisions of the Code of Federal Regulations presuppose that “technical data” first produced pursuant to a contract with a federal agency are copyrightable. 1 Similarly, federal law provides copyright protection for handbooks of standardized scientific and technical research data prepared by the Secretary of Commerce. 2 Because many of the institutions responsible for sponsoring research and for publishing research results assume that data are copyrightable, scientists may also assume that copyright provides sufficient protection. If this common belief that research data are copyrightable is inaccurate, the ramifications would profoundly affect the scientific community.

Scientists are concerned with protecting research data from competitors, especially prior to publication. Researchers may want to preserve claims to priority in discoveries, and retain the first opportunity

© 1987 High Technology Law Journal


to theorize about the data.\textsuperscript{3} They may also want to control and profit from commercial exploitation of their discoveries, or use the raw data to substantiate the quantity and quality of their work when applying for research grants. Thus, copyright protection is attractive to scientists because it secures protection from the time the work is recorded, and it gives the author exclusive rights to make copies and produce derivative works.\textsuperscript{4} Copyright in their research data would give scientists the right to reproduce and distribute their material publicly, as well as the right to make derivative works, such as articles summarizing their results. Because of this apparent protection, scientists may depend on copyright in the intangible research data.\textsuperscript{5}

Similarly, employers of scientists may have expectations that the "work-for-hire" provision of copyright law will give them control over employees' data.\textsuperscript{6} If this reliance on copyright is misplaced, research data collected in both private and academic research programs may be unprotected. Therefore, it is important to examine the exact basis in statutes and case law for the claim that research data are protectable as property.

The focus for this paper is on "raw" scientific research data, defined here to include any unedited recording of information concerning measurable properties of physical objects resulting from experimentation or controlled observation.\textsuperscript{7} This does not include charts, graphs, tables, and models which express data in summary form.\textsuperscript{8}


\textsuperscript{4} 17 U.S.C. \textsuperscript{102} (protection subsists from time work is "fixed"); 17 U.S.C. \textsuperscript{106(1)-(3)} (1982) (exclusive rights to reproduce a work, prepare derivative works, and distribute copies publicly). The definition of "derivative work" in 17 U.S.C. \textsuperscript{101} (1982) includes "abridgement, condensation, or any other form in which a work may be recast, transformed, or adapted." Preparing research data for publication in an article or presentation at a professional meeting would fall squarely within this definition since such a summary contains the "essence" of the data in detail.

\textsuperscript{5} Alternatively, scientists might rely on ownership of the physical objects that generate data to protect scientific endeavors. However, any protection provided by ownership of physical objects is beyond the scope of this Comment.

\textsuperscript{6} The work-for-hire doctrine, codified at 17 U.S.C. \textsuperscript{201(b)} (1982), grants a copyright interest in a work to the author's employer, rather than the author himself, under certain circumstances.

\textsuperscript{7} See supra notes 1 and 2 (definitions of "technical data" and "standard reference data").

\textsuperscript{8} Raw data are distinguishable from conclusions drawn from the examination of the data and from summaries of data. Conclusions involve intellectual effort rather than the mere recitation of data, and summaries involve at least a minimal transformation of the data. See infra notes 31-65 and accompanying text on the copyrightability of compilations and maps. Articles based upon research data and containing that data are copy-
The two-fold thesis of this Comment is that raw data are not copyrightable and that federal copyright law preempts the states from providing a property or quasi-property interest in raw data. Section I sets forth the basic argument for copyrightability of raw data and then counters with the problems that would result from allowing raw data to be copyrighted. In particular, Section I discusses the potential inclusion of raw data in the statutory subject matter categories, the requirement of originality as applied to factual works, the merger of idea and expression in raw data, the issue of dissemination of information versus control, and fair use of raw data. In Section II, potential forms of state protection of raw data are considered. As will become apparent, the developers of copyright case law, as well as the federal statute, did not have scientific research data in mind.

I. COPYRIGHT

The basic requirement for copyrightability under the Copyright Act of 1976 is that protection only extends to "original works of authorship fixed in any tangible medium of expression." Thus, the argument for the copyrightability of scientific research data is simple. Recorded data are "fixed" in a "tangible medium of expression" when they are written down or photographed. Furthermore, the fixed raw data constitute a "writing" within the scope of copyright because fixation is either a "literary work" under section 102(a) or a "compilation" under section 103(a). The required originality of authorship is supplied by the creativity and labor of scientists in devising experiments and collecting the data the experiments generate. In short, the argument is that raw data are writings that result from originality and thus satisfy the requirements for copyrightability. One could argue further that the constitutional objective of promoting the "Progress of Science . . . by securing for limited Times to Authors . . . the exclusive Right to their . . . Writings" favors granting copyright protection to raw data. No underlying "idea, procedure, process, system, method of operation, concept,
principle, or discovery" can be protected by copyright—protection can be claimed only for a particular "expression" that the author has produced. Even if raw data is copyrightable, other scientists would still be free to duplicate the data by conducting their own experiments. Thus, no "principle or discovery" would be protected by giving copyright protection to raw data.

A. Problems with Potential Subject Matter Categories

Unfortunately, there are various problems with the argument that raw data are copyrightable. The first problem is determining in which category of copyrightable subject matter, listed in section 102(a) of the Copyright Act, raw data belong. The categories in section 102(a) do not neatly encompass the arrays of numbers which constitute most scientific data. For example, one category that might be applicable is "literary works." However, despite the fact that the congressional reports include compilations of data within this category, scientific research data are less self-evidently "literary works" than most other factual works. Neither Congress nor the courts have addressed the particular issue of how to categorize scientific research data.

The following discussion examines three categories of copyrightable subject matter. The first category is photographs. The second and third are two broad categories of fact-gathering works—collections of facts, and non-fictional narratives. However, there are significant problems

16. Id.
with each potential category. Both collections of facts and non-fictional narratives raise the issue of whether to protect only a researcher's original contribution or, because of the vast effort expended, all the facts the researcher uncovers.

1. **Photographs**

Scientifically significant photographs such as those of astronomical events or Wilson cloud-chamber events appear to fall within the meaning of the statute.\(^{18}\) Photographs of nature are copyrightable.\(^{19}\) However, the case law on photographs presumes they are the result of artistic creativity. None of these cases involve photographs produced strictly for their informational content.\(^{20}\) Courts in photograph cases emphasize the artistic element of selecting and arranging the objects to be photographed.\(^{21}\) Such creative decisions make a photograph an original work of authorship. It can be argued that the reason the requirement of creativity is greater here than in other areas of copyright protection is that without some creativity a photograph is not an "original work" of an "author" but merely a mechanical reproduction of material in the public domain.\(^{22}\)

Some of the concerns expressed in early cases about granting copyright to photographs are relevant in the scientific setting because scientific photographs are media for recording data, rather than works of art. In one early photograph case, upholding the copyright on a posed photograph, the Supreme Court left open the issue of whether an "ordinary production of a photograph" (i.e., one involving no arrangement of the subject) is copyrightable.\(^{23}\) Such photography "is merely mechanical, with no place for novelty, invention, or originality. It is simply the

---


\(^{19}\) Cleland v. Thayer, 121 F. 71, 72 (8th Cir. 1903). The photograph in question, of Colorado scenery, was "artistically colored," and "used various original, ingenious, and artistic ideas" (e.g., arranging light and shadow).


\(^{21}\) See, e.g., Bleistein v. Donaldson Lithographic Co., 188 U.S. 239, 250 (1903) (the least pretentious picture has more originality in it than directories which may be copyrighted); Burrow-Giles Lithographic Co. v. Sarony, 111 U.S. 53, 60 (1884); Jewelers' Circular Publishing Co. v. Keystone Publishing Co., 274 F. 932, 934 (S.D.N.Y. 1921), *aff'd*, 281 F. 83 (2d Cir.), *cert. denied*, 259 U.S. 581 (1922) (no photograph, however simple, can be unaffected by the personal influence of the author, and no two will be absolutely alike).

\(^{22}\) See Note, "Expression" and "Originality" in Copyright Law, 11 Washburn L.J. 400, 404 (1972).

manual operation, by the use of these instruments and preparations, of transferring to the plate the visible representation of some existing object, the accuracy of this representation being its highest merit."²⁴

Blindly snapping a camera lens is precisely what occurs in scientific experiments. Scientists do select the area of study and prepare experiments, but they exercise no further control over what is revealed by their experiments. The scientists must, in the sense important to the issue at hand, blindly snap the shutter once the topic of the experiment is prepared for the experiment to be valid. If scientists were more manipulative, their experimental results would reflect only their theories and would not be "objective."²⁵ Thus, research photographs do not fit within the Supreme Court's characterization of a photograph as "the personal reaction of an individual upon nature."²⁶

With regard to scientific photographs, the issue is whether the originality requirement of copyright can be satisfied by the preparation and design of experiments.²⁷ Research photographs are the product of "creative intellectual or aesthetic labor."²⁸ But it is only creativity of expression, rather than labor, which is relevant to copyright law. Research photographs, like research data in general, do not contain creativity of expression.²⁹

Photograph cases under the Copyright Act offer no guidance on whether the resulting data embodied in the picture are copyrightable. Because research photographs are devices for recording data and are not intended to be works of art,³⁰ it may make more sense to treat scientific research photographs, along with computer printouts and handwritten arrays of data, as "fact-gathering" works.


²⁶ Bleistein, 188 U.S. at 250.

²⁷ Concerning the photograph of a street scene, the court in Pagano v. Beseler Co. said: "[i]t undoubtedly requires originality to determine just when to take the photograph, so as to bring out the proper setting for both animate and inanimate objects, with the adjunctive features of light, shade, position, etc." 234 F. 963, 964 (S.D.N.Y. 1916).


²⁹ See also infra notes 53-56 and accompanying text on the selection and arrangement rationale.

³⁰ In general, scientists are only interested in the factual context of their photographs. However, scientists may also want to copyright research photographs as art in order to exploit them commercially (e.g. in the form of posters or books). This raises problems different from those discussed in this Comment.
2. Compilations and Maps

Although the general rule is that facts are not protected by copyright, certain types of fact-gathering works, such as compilations and maps, are protectable. The 1976 Act defines compilations as works "formed by the collection and assembly of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship." Maps are pertinent to research data issues because courts sometimes require actual field observations to support claims for copyright protection. However, while maps are technically within the category of "pictorial, graphic, and sculptural" works, they can be analyzed as compilations because they involve selection, synthesis, and judgment like other fact-gathering works. Accordingly, no separate discussion is required.

There are two distinct lines of cases that ascribe different degrees of protection to compilations of facts. The first line of cases follows the literal words of the 1976 Act, which states that: "[t]he 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, and does not imply any exclusive right in the preexisting material." Thus, under this line of cases, it is the selection and arrangement of material by the researcher which, if more than trivial, is the


33. For example, one court held that the requirement of a "modicum of creative work" is satisfied only where the "publisher of the map in question obtains some of that information by the sweat of his own brow." Amsterdam v. Triangle Publications, 189 F.2d 104, 106 (3d Cir. 1951).


protectable element in a collection of preexisting material. As a result, the facts gathered can be used by any person who does not also take the researcher’s intellectual efforts of selecting those particular facts and arranging them in that particular order.

The second line of compilation cases rewards the labor of the researcher by giving him a property interest in the material gathered. This material is protected from substantial copying, although anyone is free to duplicate the research provided she does not copy the prior results. In other words, in this line of cases protection for labor expended extends to the content of fact-gathering works (the facts themselves) rather than merely to the author’s expression (selection and arrangement).

However, the “rationale behind protecting such compilations . . . is not clear” since copyright in compilations is supposed to protect only original material contributed by an author. One court explained the rationale for such extensive protection as follows:

The compiler’s contribution to knowledge normally is the collection of the information, not its arrangement. If his protection is limited solely to the form of expression, the economic incentives underlying


39. See, e.g., Eckes v. Card Prices Update, 736 F.2d 859, 862 (2d Cir. 1984). Compilations are protected “regardless of whether the individual items in the material have been or ever could have been subject to copyright.” USCCAN, supra note 15, at 57. Therefore, even if each scientific datum is uncopyrightable because it is a fact, the compilation may still be copyrightable. See infra notes 79-112 and accompanying text. Protecting a compilation when the compiler selects and arranges from items already assembled by another, rather than also initially assembling the items, is open to the criticism that it emphasizes selection to the exclusion of the “collection and assembly” element in the definition of “compilations.” See Ginsburg, Fact Works Revisited, 192 N.Y.L.J. 22 (1984); Patry, supra note 32, at 14. The raw data resulting from scientific experimentation fulfills this “collection” requirement.


41. See, e.g., Schroeder, 566 F.2d at 5; Leon, 91 F.2d at 486; Jeweler’s Circular, 281 F. at 88 (a compiler of a directory produces by his labor a meritorious composition in which he may obtain a copyright); Rand McNally, 591 F. Supp. at 736; Financial Information, 599 F. Supp. at 999 n.7 (copyright protection should be afforded the compiler’s industry even when his arrangement is not copied because often that is the only way that protection will prove meaningful); National Business Lists, 552 F. Supp. at 94; Southwestern Bell, 371 F. Supp. at 906; Triangle Publications v. New England Newspaper Publishing Co., 46 F. Supp. at 198. Cf. Dow Jones, 546 F. Supp. at 115.

42. Rand McNally, 591 F. Supp. at 731.

43. Yet the court in Rand McNally followed Seventh Circuit precedents and granted copyright protection to “industrious collections.” Id. at 731-32.
the copyright laws are largely swept away. Recognizing this, the courts have long afforded protection under the copyright laws against appropriation of the fruits of the compiler’s industry.\textsuperscript{44} However, that court recognized that such “protection does not fit nicely into the conceptual framework of copyright law and has for that reason been criticized.”\textsuperscript{45} Indeed, granting protection on the basis of the labor expended in making a compilation has been criticized by Professors Nimmer and Gorman for expanding copyright protection by importing ideas into the field of copyright which are unrelated to the protection of expression.\textsuperscript{46}

Recently, courts have rejected the “sweat of the brow” rationale in both a classic factual compilation case and a classic map case.\textsuperscript{47} In the former, the Court of Appeals for the Second Circuit recently stated: “[t]o grant copyright protection based merely on the ‘sweat of the author’s brow’ would risk putting large areas of factual research material off limits and threaten the public’s unrestrained access to information.”\textsuperscript{48} In the latter, the court explicitly rejected the sweat of the brow requirement stating that authorship for maps exists in “selection, design, and synthesis.”\textsuperscript{49}

Moreover, scientific research data do not fit either the “selection and arrangement” or the sweat of the brow rationale well. Consider first the selection and arrangement rationale. Raw data consists of very large collections of informational items. Attempting to characterize raw data as a “compilation” is problematic, however, because the order of raw data is dictated by the laws of nature, not by the creativity of the scientist.\textsuperscript{50} A scientist does not synthesize data, he only records it.\textsuperscript{51} The arrangement of data is predetermined as intractably as is the

\begin{itemize}
\item\textsuperscript{44} National Business Lists, 552 F. Supp. at 92.
\item\textsuperscript{45} \textit{Id.} See \textit{supra} text accompanying note 32 (definition of compilation).
\item\textsuperscript{46} 1 M. NIMMER, THE LAW OF COPYRIGHT § 3.04, at 3-19 to -20 (1983); Implications for Copyright, \textit{supra} note 17, at 572. But see, Denicola, \textit{supra} note 23, at 519-24; Hill, \textit{supra} note 17. See also, National Business Lists, 552 F. Supp. at 95 (directory cases illustrate that misappropriation doctrine has long found a house, if not a home, in copyright law and that notions of unfair competition are implicit in copyright protection of diligent application).
\item\textsuperscript{47} Financial Information v. Moody’s Investor Serv., 808 F.2d 204 (2d Cir. 1986) (compilation); United States v. Hamilton, 583 F.2d 448 (9th Cir. 1978) (map).
\item\textsuperscript{48} Financial Information v. Moody’s Investor Serv., 808 F.2d at 207. See also Triangle Publications v. Sports Eye, Inc., 415 F. Supp. at 685-86 (copyright protection extends only to the method or form for expressing the data).
\item\textsuperscript{49} Hamilton, 583 F.2d at 452.
\item\textsuperscript{50} Because data are generated in an automatic manner, if each datum is not individually copyrightable, no quantity of data will be copyrightable. See \textit{infra} notes 90-112 and accompanying text.
\item\textsuperscript{51} Even maps which disclose previously unknown facts involve skill and discretion in the selection process. Nothing similar occurs in the production of scientific data.
\end{itemize}
chronological ordering of historical events. The only scientifically significant order is one which illustrates the law of nature or other phenomenon discovered by the scientist. There is no room for variation in the presentation and thus none for individual choice.\(^5\)

The higher the degree of selectivity and judgment in creating a compilation, the stronger the claim is to copyright protection. However, collection of raw data involves little creativity. From the point of view of copyright law, a scientist’s only contribution to the data collected is the exercise of creativity and effort in determining which phenomena to study and which variables to correlate.\(^5\) This effort amounts to no more than choosing a subject within the public domain to copy and determining a framework for copying. As Judge Wyzanski put the point: “[t]o constitute a copyrightable compilation, a compendium must ordinarily result from the labor of assembling, connecting, and categorizing disparate facts which in nature occurred in isolation. A compilation, in short, is a synthesis.”\(^5\)

Yet it is not always clear whether scientific experimental and observational data can be characterized as an analysis of a series of events rather than an analysis of a single, larger event. Often research data consist only of a picture or record of a single occurrence. Such data do not fit easily into the subject matter category of fact compilations. As Judge Wyzanski stated: “[i]t is rare indeed that an analysis of any one actual occurrence should be regarded as a compilation.”\(^5\)

There is no copyrightable element of transformation or synthesis of the public domain material recorded during a scientific experiment or observation even if there is creativity in designing the experiment. Although, “practically anything novel can be copyrighted,” some novelty in expression is needed.\(^5\) As a result, research data are not copyrightable as fact compilations under the selection and arrangement criteria.

---

52. The order in which data from different experiments is reported is open to variation but is scientifically insignificant if chronological order is not involved; if chronological order is relevant, the order is dictated by the data.

53. The simple correlation of names and telephone numbers was held to be a copyrightable element of a directory in New York Times Co. v. Roxbury Data Interface, 434 F. Supp. 217 (D.N.J. 1977). The selection of variables in science may involve comparable creativity, and thus support copyrightability of research data in some jurisdictions. However, no other court has placed strong emphasis on correlation of variables. Nor has any court granted copyright protection solely based on the existence of a correlation pattern, without evidence of significant effort, selection, or arrangement.


55. Id. Accord Financial Information, 599 F. Supp. at 997 (assembling a handful of statistics about a single, solitary occurrence not copyrightable).

At first glance the sweat of the brow rationale seems to provide a better basis for copyright protection for scientific data. The standard of originality is not especially high. For example, under the sweat of the brow rationale, telephone books have been given protection for merely placing names submitted to the telephone company in conventional alphabetical order.57 As Judge Learned Hand stated: ""[t]he 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 numbers, acquires material of which he is the author."58 In addition, computer data-bases are copyrightable59 although it makes little sense to speak of an "order" in these automated compilations.

One can argue that scientists should be given copyright protection for data resulting from their experiments, which are often carried out at great time and expense, because compilations "have value because the compiler has collected data which otherwise would not be available."60 Even where the order of data is dictated by nature, scientists frequently reveal information not otherwise available to the public. In addition, giving protection would force other scientists to repeat experiments, thereby checking their predecessors' results. One can argue that the practice of checking previous results by replication would lead to more reliable theories, and should be encouraged.

Unfortunately, under current practice repeating experiments to test their validity "is a myth, a theoretical construct dreamed up by the philosophers and sociologists of science."61 Grants are not normally awarded for checking other scientists' results. Scientists repeat experiments only if the earlier results are controversial; otherwise, scientists accept the findings and build upon them.62 Thus, one could argue that an incentive to duplicate research is needed.

However, the reasons enunciated for protecting works of diligence do not carry over to the domain of scientific research. The "sweat of the brow" rationale is invoked only when it is necessary to provide the incentive of a property interest for a socially valuable but otherwise

57. It has been suggested that this feature makes telephone books the outer limit of a copyrightable compilation. Latman & Ginsburg, supra note 40, at 2, col. 1.
58. Jeweler's Circular, 281 F. at 88. This approach finds "authorship in the act of aggregating isolated pieces of information." Denicola, supra note 17, at 530.
59. 17 U.S.C. at § 101-102 (1982); USCCAN, supra note 15, at 54. Stating a set of facts in a form that is computer-usable may not involve the authorship necessary for copyright protection. But most data-bases, unlike scientific experiments which produce data, also involve the selection of data.
62. See id. at 60-87.
unprotectable work resulting from mere diligence and tedious labor. Although scientists must pay great attention to exacting detail, scientific experiments involve great creativity in their design and in their interaction with theory. Thus, scientific experiments involve more imagination and inventiveness than do works of mere diligence. Scientists attempt to find new phenomena and develop new theories for understanding nature. In this sense, the research process is more of an "art" than is fact-gathering. Conducting experiments is not an end in itself but only one component of a creative enterprise; a separate incentive for the production of research data is not required. Especially since the legitimacy of the "sweat of the brow" rationale in copyright law has been questioned, this rationale should not be expanded into a creative field such as scientific research.

Furthermore, scientific research differs from other fact-gathering activities in that the resulting data cannot be used piecemeal, whereas facts in directories or parts of a map can be so used. It is the whole of the data that is important to scientists. An individual datum revealed by an experiment is usually of little use apart from the pattern disclosed. A useful directory can either be limited to selected highlights of a subject or can be comprehensive. On the other hand, enough scientific data on a particular subject must be produced to support claims concerning alleged discoveries, although not all instances of the alleged phenomena need be produced. In other words, scientific data must be comprehensive, but not exhaustive. Thus, the protection copyright gives against use of even a portion of a copyrighted work is not necessary.

In short, a collection of scientific data does not easily fit the definition of a "compilation" within the meaning of the 1976 Act. The "sweat of the brow" approach has been used in cases involving fact-gathering activity which are significantly different from scientific research. Scientists simply do not string together facts in the manner contemplated by courts in providing protection for compilations.

---

63. See supra text accompanying note 17 for discussion of criticism of the "sweat of the brow" rationale. As noted, some current cases and scholarly studies support the rationale.

64. Directories are also end-products in themselves; raw data are the subject matter for further scientific study insofar as experiments and theories interact with each other in the scientific process.

65. See supra text accompanying note 32 (defining "compilation").
3. Non-Fictional Narrative Works

The third branch of fact-gathering cases involves narrative works of fact: news, biography, and history. There are two schools of thought about whether the research supporting such works can be copyrighted. The less popular school of thought protects the fruit of a researcher's labor by requiring that subsequent authors do independent research from the original sources and that they not make "substantial and unfair use" of the first researcher's work. The justification given is that the "substantial investment of time, money, and labor" expended in researching for a work should be protected from appropriators.

For example, the district court in Miller v. Universal City Studios held that "[t]he law is clear that research can be copyrightable." That court viewed "the labor and expense of the research involved in . . . obtaining . . . those uncopyrightable facts to be intellectually distinct from those facts and more similar to the expression of the facts than to the facts themselves." On appeal, however, the holding that research is copyrightable was reversed. The appellate decision typifies the second school of thought, which is the majority view regarding the copyrightability of research.

The valuable distinction in copyright law between facts and the expression of facts cannot be maintained if research is held to be copyrightable. There is no rational basis for distinguishing between facts and the research involved in obtaining facts. To hold that research is copyrightable is no more or no less than to hold that the facts discovered as a result of research are entitled to copyright protection. "One who narrates matters of fact may be protected by copyright as to his arrangement, manner and style, but not as to material of ideas therein set forth." Oliver v. Saint Germain Found., 41 F. Supp. 296, 299 (S.D. Cal. 1941).

See also MCA, Inc. v. Wilson, 677 F.2d 180, 183 (2d Cir. 1981) (dictum) (use of copyrighted material without owner's consent generally not considered reasonable if it extensively copies or paraphrases the original or bodily appropriates the research upon which the original was based); Eisenschiml v. Fawcett Publications, 246 F.2d 598 (7th Cir.), cert. denied, 355 U.S. 907 (1957); Holdredge v. Knight Publishing Corp., 214 F. Supp. 921, 922-23 (S.D. Cal. 1963); Huie v. Nat'l Broadcasting Co., 184 F. Supp. 198, 200 (S.D.N.Y. 1960) (dictum) (publishing a history re-written from another historian's book without any independent research constitutes infringement).

Miller, 460 F. Supp. 984, 987 (D.C. Fla. 1978) (citations omitted) (arguing that this protection rewards individual effort and ingenuity in obtaining knowledge), rev'd, 650 F.2d 1365, 1372 (5th Cir. 1981).

Miller v. Universal City Studios, 650 F.2d 1365, 1372 (5th Cir. July 1981).
Similarly, in *Harper & Row v. Nation Enterprises* Justice Brennan, in his dissent, discussed whether the use of factual material from a manuscript describing particular historical events (but not the direct copying of the manuscript) infringed the author’s copyright in the manuscript. Addressing the issue of whether facts are copyrightable, Brennan noted that “[w]ere an author able to prevent subsequent authors from using . . . facts contained in his or her work, the creative process would wither and scholars would be forced into unproductive replication of the research of their predecessors.” Brennan went on to find that, because “[a]part from the quotations, virtually all of the material in [the allegedly infringing article] indirectly recounted [the plaintiff-author’s] factual narrative, . . . [n]o copyright can be claimed in this information *qua* information.”

Following Justice Brennan’s reasoning, allowing copyright in scientific data would amount to allowing copyright in “information *qua* information,” which would stifle creative research and promote wasteful duplication of scientific effort. Thus, this school of thought rejects copyright protection for research, and instead emphasizes the factual nature of research and the waste involved in duplicating the effort of the first researcher.

district court and court of appeals agreed that “[a]s was the case with ideas, if the expression arrangement and selection of the facts must necessarily, by the nature of the facts, be formulated in given ways then they are not copyrightable.” *Id.* at 1368. As discussed above, this claim is directly applicable to the situation involving scientific research data. See supra note 52c and accompanying text. The court of appeals also stated that “[a] copyright in a directory however, 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.” *Miller*, 650 F.2d at 1369. Thus, this court also rejected the “sweat of the brow” approach to copyright protection.

73. The majority had decided that the quotation of approximately 300 words from the manuscript was not “fair use” and therefore did not reach the question of copyrightability of the historical facts. The majority did note, however, that “[n]o author may copyright his ideas or the facts he narrates.” *Id.* at 556 (dictum).
74. *Id.* at 582 (Brennan, J., dissenting).
75. *Id.* at 583.
76. See also *Hoehling v. Universal City Studios*, 618 F.2d 972 (2d Cir. 1980); *Rosemont Enters. v. Random House*, Inc., 366 F.2d 303 (2d Cir. 1966); *Suid v. Newsweek Magazine*, 503 F. Supp. 146 (D.D.C. 1980); *Marshall v. Yates*, 27 *PAT. TRADEMARK & COPYRIGHT J.* (BNA) 137 (C.D. Cal. 1983). The second circuit has “clearly repudiated Toksvig and its progeny” on the copyrightability of research. *Hoehling*, 618 F.2d at 979; but see *MCA, Inc. v. Wilson*, 677 F.2d at 180, 183 (2d Cir. 1981). In a comment on *Miller*, it has been suggested that the court’s “holding that research is not copyrightable may deny protection to some works in which research is the only original element.” *Sato*, *Copyright Law and Factual Works—Is Research Protected?*, 58 WASH. L. REV. 619, 627 (1983).
Scientific writings pose the same problems as nonfictional narratives with regard to the copyrightability of the underlying research. First, like research for other nonfictional narrative works, scientific research represents the expenditure of significant time and labor. Second, protection of raw data would force subsequent researchers to duplicate the effort of a prior researcher. Currently, when a scientist publishes an article containing theories and summaries of data, the article as a whole is copyrightable. Other scientists can criticize the results by either conducting new experiments in order to produce data indicating a position contrary to that advocated in the article, or by re-analyzing the published data to reveal an error in the original analysis. If the raw data contained in the article were copyrightable, the latter option would be unavailable.

The conflict between the two schools of thought on the copyrightability of research should be resolved in favor of the majority view precluding copyright protection for raw data, at least where scientific research is involved. Granting protection would force subsequent researchers to repeat experiments. While such duplication would serve the useful purpose of checking earlier results, it would frequently result in a wasted effort. Furthermore, scientific data collection involves less of a selection process than historical or biographical works, as scientific data are collected comprehensively rather than selectively from the available data. Thus, the rationale for protecting the research effort itself is weaker for scientific data than for other nonfictional works.

To summarize, the problem of placing raw data in a category of copyrightable subject matter is formidable. Raw data are too integral a part of a process involving too much imagination to justify invoking the "sweat of the brow" rationale, yet are too rigidly dictated by nature to justify the "selection and arrangement" rationale. Therefore, raw data fall outside the recognized categories of copyrightable subject matter.

B. Authorship & Originality

Copyright protects only an author's original contribution. In many ways, this requirement is extremely lax. One court has said: "[a]ll

77. See supra note 61 and accompanying text.
78. See supra note 53 and accompanying text.
79. Section 102(a) of the 1976 Act extends copyright protection only to "original works of authorship." The 1909 Act had extended protection to "all the writings of an author." 17 U.S.C. § 4 (1909) (repealed 1976). "Originality" is left undefined in the 1976 Act, but it is doubtful that if scientific data are not protectable under the 1976 Act they would have been protectable under the earlier Act since the phrase "original works of authorship" is "intended to incorporate without change the standard of originality established by the courts under the present [1909] statute." USCCAN, supra note 15, at 51.
that is needed to satisfy both the Constitution and the statute is that the ‘author’ contributed something more than a ‘merely trivial’ variation, something that can be recognized as ‘his own.’” 80 Originality in this context “means little more than a prohibition of actual copying.” 81 However, if scientists merely record public domain material, they contribute nothing to the expression; if nothing has been added by an individual researcher, it makes no sense to speak of an “author” or “originality.”

The Copyright Office Regulations promulgated under the 1909 Copyright Act (but still applicable under the 1976 Act) deny copyright to “[w]orks consisting entirely of information that is common property containing no original authorship, such as, for example: standard calendars, height and weight charts, tape measures and rulers, schedules of sporting events, and lists of tables taken from public documents or other common sources.” 82 Like the works listed in this regulation, raw scientific research data do not contain sufficient expression and originality to be protected by copyright law.

1. Originality of Expression

A compiler of disconnected facts makes an original contribution by the selection and arrangement process or by the labor expended in collecting the material. 83 Such a compiler is an “author” of a writing, i.e., one “to whom anything owes its origin; originator; maker; one who completes a work of science or literature.” 84

An axiom of copyright law is that the Act protects only the expressions of ideas, not the ideas themselves. 85 Thus, “[t]here is no copyright of facts,” 86 as no one may claim original expression in facts. 87 When any expression is so “straightforward and simple” as virtually to “spring

81. See generally Olson, Copyright Originality, 48 Mo. L. Rev. 29 (1983).
83. See supra notes 31-42 and accompanying text.
86. Greenbie v. Noble, 151 F. Supp. 45, 66 (S.D.N.Y. 1957). See also supra note 31 and accompanying text; USCCAN, supra note 15, at 56 (copyright does not preclude others from using the ideas or information revealed by the author’s work).
directly' from uncopyrightable material, there is no "original creative authorship." 88

The court in Alfred Bell & Co. v. Catalda Fine Arts 89 proposed a test for determining when a work based on public domain material will support a copyright. The court stated that "a 'copy of something in the public domain' will support a copyright if it is a 'distinguishable variation.'" 90 Scientific research data fail this distinguishable variation test. The facts expressed by the raw data are stated in the simplest language possible: scientific notation. The barest description of an event is, in the eyes of copyright law, a writing without an author.

Although facts are sometimes discovered by an author, they are not themselves works of authorship. As Professor Nimmer remarked, "[o]ne who discovers an otherwise unknown fact may well have performed a socially useful function, but the discovery as such does not render him an 'author' in either the constitutional or statutory sense." 91 Authorship requires originality of expression, not merely the discovery of a fact.

The fifth circuit has also expressed the opinion that facts are akin to discoveries, not original works:

Obviously, a fact does not originate with the author of a book describing the fact. Neither does it originate with one who "discovers" the fact. "The discoverer merely finds and records. He may not claim that the facts are 'original' with him although there may be originality and hence authorship in the manner of reporting, i.e., the 'expression,' of the facts." 92

A "discovery" has been judicially defined as the "disclosure of an hitherto unknown fact, principle, or theory." 93 Such discoveries are the substance of the work of scientists, but discoveries, along with ideas,

88. Morrissey v. Proctor Gamble Co., 379 F.2d 675, 679 (1st Cir. 1967). The scope of copyright protection increases with the extent that expressions can differ from the underlying idea. Sid & Marty Krofft Television Prods. v. McDonald's Corp., 562 F.2d at 1168. At the margin where expression and idea merge, there is no protection and so the work is uncopyrightable. To claim that such a work is copyrightable but that there is no protection provided against copying is conceptually wrong: it makes no sense in copyright analysis to say that a work is copyrightable but has no protection.

89. 191 F.2d 99, 102 (2d Cir. 1951).

90. Id. (footnote omitted). Ideas and facts are not "writings" under Article I, section 8, clause 8 of the Constitution, even though courts give "writings" an expansive interpretation. See Goldstein v. California, 412 U.S. at 561; Rubin v. Boston Magazine Co., 645 F.2d 80, 83 (1st Cir. 1981). The Supreme Court in Burrow-Giles, 111 U.S. at 58, defines "writings" for the Copyright Clause to embrace "all forms . . . by which the ideas in the mind of the author are given visible expression." However, the bare statements of facts are too minimal an "expression" to qualify.

91. 1 M. Nimmer, supra note 46, §§ 2.11[A] and 2.11[E], at 2-158 and 2-168.


procedures, and processes, are not copyrightable "regardless of the form in which [they are] described, explained, illustrated, or embodied" in an original work of authorship. Accordingly, scientific data do not meet the requirement that copyright subject matter be an expression owing its "origin" to an author.

2. Facts as Expressions of Theory

Post-empiricist philosophers believe that a scientist mixes theory with observation to generate facts in such a manner that all data are "expressions" of an underlying theory. In other words, there are no "bare facts" since every fact is "an event as we see it" reflecting the theory of a particular scientist. The copyright argument is thus that a scientist's world view influences what she deems "facts" and that therefore, her observation and record of data creates a tangible "expression" of her ideas and theories.

However, even if raw data are theory-laden, the data present the facts from one point of view in a very simple manner. Scientists select questions to answer and experiments to perform, thereby preselecting the type of data that will result and providing in advance a framework with which to interpret the data. This preselection does not mean that the recorded data have any additional "expression" in them. Every fact requires some conceptualization, but raw data includes no expression apart from this bare conceptualization.

94. 17 U.S.C. § 102(b) (1982). This principle simply makes clear that ideas are absolutely not protectable. Thus, while discoveries are not copyrightable, an original expression describing one would be. Rubin, 645 F.2d at 82.

95. N. HANSON, OBSERVATION & EXPLANATION: A GUIDE TO PHILOSOPHY OF SCIENCE 1, 9 (1971).

96. One author argues that research is more than the "mindless collection of facts." He asserts that financial research is analysis and interpretation of events in addition to the organization the author imposes on the raw facts. Note, Copyright Law—Will the Denial of Copyright to an Author's Research Impede Scholarship? Miller v. Universal City Studios, Inc., 605 F.2d 1365 (5th Cir. Jul. 1981), 5 W. NEW ENG. L. REV. 103, 116-17 (1982).

97. S. LANGER, PHILOSOPHY IN A NEW KEY 227 (1951).

98. N. HANSON, supra note 95, at 11-12.

99. For each fact, some conceptual element is required. It makes no sense to speak of "unexpressed facts" or of informational content apart from expression. For example, the sun exists apart from expressions of it, but our use of language requires us to speak of facts in expressive language, such as, "the sun exists" or "the sun is a star." In this sense, facts are always "expressions."
To produce a copyrightable expression describing a single event, the account of the event must either have individuality of expression or reflect the author's "peculiar skill and judgment." The bare reporting of events that occurs in scientific experimentation lacks such skill and judgment. Designing an experiment preselects certain data for expression, but this is comparable only to choosing which public domain fact to copy—nothing separately copyrightable is involved. There might be more complicated ways to express simple facts: for example, Einstein could copyright books written in a standard language explaining the significance of the formula "E = mc^2" but the formula itself is not copyrightable. Likewise, the simplest expression in scientific language of any idea or fact is not copyrightable, although a complicated expression would be.

Each scientific datum gives a simple (usually mathematical) description of one fact. In addition to each individual datum describing one fact, the data collectively "merge" into a single fact which the observer can understand as an automatic expression of a natural law or pattern. The expression embodied in the data is inseparable from the underlying fact. In particular, one court has treated such expression as uncopyrightable because, "[c]opyright protection will not be given to a form of expression necessarily dictated by the underlying subject matter." Similarly, in a historical research case a court held that "if the expression arrangement and selection of the facts must necessarily, by the nature of the facts, be formulated in given ways then they are not copyrightable.""104

101. See 17 U.S.C. § 102(b) (providing that no "principle" or "discovery" can be copyrighted). See also Miller v. Universal City Studios, Inc., 650 F.2d 1365 (5th Cir. July 1981); Rubin v. Boston Magazine Co., 645 F.2d 80 (1st Cir. 1981).
102. See 1 M. Nimmer, supra note 46, § 2.11[A] (data collectively "express" one fact or idea, and this itself is another fact). Note that protectable compilations as a class are not exceptions to the merger doctrine if the "sweat of the brow" rationale is rejected in favor of the "selection, and arrangement" rationale. See supra notes 31-49 and accompanying text. Different "expressions" are then possible from the same facts.
104. Miller v. Universal City Studios, 650 F.2d at 1368 (part of jury instruction not challenged on appeal). See also 1 M. Nimmer, supra note 46, § 2.11[B] (no protection will be accorded to the literal form of expression of a fact if such form does not evidence originality); see also Nichols v. Universal Pictures Corp., 45 F.2d 119, 121 (2d Cir. 1930) (Learned Hand, J.) (distinguishing ideas and expression in a play). Justice Brennan cited Nichols to support the notion that copyright law must proscribe more than literal appropriation of the author's work. Otherwise, "a plagiarist could avoid infringement by immaterial variations." Harper & Row, Publishers v. Nation Enters., 471 U.S. 539, 583 n.5 (1985) (Brennan, J., dissenting).
One important corollary to the idea/expression distinction\(^{105}\) is that an expression will receive copyright protection only if it is possible to create alternative expressions of the same idea involving substantial variation.\(^ {106}\) An expression of uncopyrightable subject matter is itself uncopyrightable if only a limited number of ways to express the given subject matter are available.\(^ {107}\) If any one expression were given copyright protection, a virtual monopoly over an idea or fact would result.\(^ {108}\)

Therefore, because the primary concern with the free flow of ideas prevails over the property interest, the law denies copyright protection.

When the "idea" and its "expression" are thus inseparable, copying the "expression" will not be barred, since protecting the "expression" in such circumstances would confer a monopoly of the "idea" upon the copyright owner free of the conditions and limitations imposed by patent law.\(^ {109}\)

Thus, where idea and expression merge, neither is protected by copyright law.

Scientific research data are the paradigm of the merger of idea (or fact) and expression—"the information expressed is the data. Idea and expression coincide when "the expression provides nothing new or additional over the idea,"\(^ {110}\) and this is precisely what occurs with scientific

---

105. See supra notes 85-94 and accompanying text (discussing idea/expression distinction).


107. Id. at 678-79. See also Landsberg v. Scrabble Crossword Game Players, Inc., 736 F.2d 485, 488-89 (9th Cir. 1984), cert. denied, 469 U.S. 1037, appeal after remand, 802 F.2d 1193 (9th Cir. 1986) (infringement of factual works).

The same is true in the area of art: if a form is dictated solely from functional considerations or if there is no room for significant variation, the court will not grant copyright protection. See, e.g., Herbert Rosenthal Jewelry Corp. v. Kalpakian, 446 F.2d 738 (9th Cir. 1971); Trifari, Krussman & Fishel, Inc. v. Charel Co., 134 F. Supp. 551 (S.D.N.Y. 1955).

Fact-gathering photographs may be different because research photographs are usually not the only expression of an event. See Time, Inc. v. Bernard Geis Assocs., 293 F. Supp. 130, 143-44 (S.D.N.Y. 1968)(protected photographs of the assassination of Kennedy). If the photographed event is also somehow uniquely important and integral to a theory, under the doctrine of Time it may be copyrightable but open to a "fair use" exception. Id. at 146. To be scientifically significant, events must have unrepeatable features (i.e., features revealing lawful aspects to events). Thus, only events which occur rarely (e.g., certain rarely observed astronomical events) would in principle qualify for such an exception if copyright protection is permitted.

108. However, arriving at the same idea or fact through completely independent research would still be allowed. See supra text accompanying notes 41-43.


110. Sid & Marty Krofft Television v. McDonald's Corp., 562 F.2d 1157, 1168 (9th Cir. 1977). The court goes on to say that "when idea and expression coincide, there will be protection against nothing other than identical copying of the work." Id. This conflicts with the notion that when idea and expression merge, "protecting the expres-
data and facts. Conjectures concerning new discoveries or theories are interpretations of the data, not something already "in" the data themselves. If the idea is taken to be the experimental procedure, then the idea is so detailed that only one pattern of data could result. In this case, the variation in expressions represented by differing sets of data would be trivial, and thus the expression would collapse into the idea. In simplest terms, the data as a whole express a law or pattern of nature which the underlying reality produces. Thus, the data collectively express one idea in addition to each datum reporting an individual fact.\textsuperscript{111} Therefore, protecting the data collectively or individually proves just as difficult as protecting the data individually because there is no protectable expression distinct from the underlying unprotectable facts.

Such data are forms of expression which cannot be varied without altering the facts expressed.\textsuperscript{112} There is no room for a plurality of expressions, and hence no room for the creativity protected by the copyright laws. Thus, based on the statutory and case law, copyright protection probably does not extend to raw scientific data. However, because the law is not entirely clear on this issue, the policy considerations which influence copyright doctrine should be examined.

C. Policy considerations: Dissemination of Information Versus Control

Copyright law has two basic goals: (1) to encourage the dissemination of information; and (2) to provide an incentive to authors by granting them property rights in their works for a limited period of time.\textsuperscript{113} Accordingly, the decision to grant or deny copyright protection to research data follows from analyzing the tension between the free flow of information and the great control which results from providing a property interest in original works of authorship. Scientists need control of their data for research and theorizing, and copyright protection could provide that control. On the other hand, society has an interest in the dissemination of research data.

\begin{footnotesize}
\begin{enumerate}
\item To argue that individual data must be copyrightable in order to protect the compilation of them is to "boot-strap" protection of the compilation. See Financial Information v. Moody's Investor Serv., 751 F.2d 501, 511 (1984) (Newman, J., concurring), district court opinion on remand aff'd, 808 F.2d 204 (2d Cir. 1986). Without an original contribution to the expression in the compilation, the "whole" is not greater than the sum of the "parts."

\item Atari v. Amusement World, 547 F. Supp. at 228 (concerning ideas).

\item See 1 M. Nimmer, supra note 46, § 1.03[A], at 1-31 to -32.2.
\end{enumerate}
\end{footnotesize}
Copyright law attempts to resolve this tension by granting protection only to "expressions" and not to the underlying "ideas."114 Society's interest in the free dissemination of ideas is promoted, while at the same time, creative endeavors are encouraged through the protection of their expression. Thus, "[t]he public interest in the free flow of information is assured by the law's refusal to recognize a valid copyright in facts."115

One may argue, however, that copyright protection of raw data would actually aid, rather than hinder, the free flow of information. In other words, the two primary goals of copyright law may be complementary rather than conflicting. For example, the Code of Federal Regulations provides that: "[i]n order to enhance the transfer or dissemination of information produced at Government expense, contractors may be permitted to establish copyright in the data first produced in the performance of work under a contract containing [a specific clause]."116 This provision evidences Congress' concern that private publishers of handbooks of standardized reference data need some incentive to publish data produced at government expense. Copyright protection makes feasible the participation of private publishers in the program by protecting a publisher's investment.117

114. The expression/idea distinction also resolves the tension between copyright and the First Amendment, i.e. the tension between restricting an individual's ability to express himself and protecting a previous author's work. See Harper & Row Publishers v. Nation Enters., 471 U.S. 539, 556 (1985); Sid & Marty Krofft Television, 562 F.2d at 1170 (the idea-expression dichotomy already serves to accommodate the competing interests of copyright and the first amendment); Pacific and Southern Co. v. Duncan, 572 F. Supp. 1186, 1192-93 (N.D. Ga. 1983); Denicola, Copyright and Free Speech: Constitutional Limitations on the Protection of Expression, 67 CALIF. L. REV. 283 (1979); Nimmer, Does Copyright Abridge the First Amendment Guarantees of Free Speech and Press?, 17 UCLA L. REV. 1180 (1970); 1 M. NIMMER, supra note 46, § 2.11[E]; Comment, Copyright and the First Amendment: Where Lies the Public Interest?, 59 TUL. L. REV. 135 (1984). On the importance of keeping scientific expression free from restrictions, see Ferguson, Scientific and Technological Expression: A Problem in First Amendment Theory, 16 HARV. C.R.-C.L. L. REV. 519 (1980) (copyright issue not considered).


117. S. REP. NO. 1230, 90th Cong., 2nd Sess. (1968), reprinted in 1968 U.S. CODE CONG. & ADMIN. NEWS 2580, 2585-86. Research data produced under government research grants independent of the program envisioned by the Standard Data Reference Program probably do not qualify as works of the United States Government under § 105 of the 1976 Act (which would disqualify the works for copyright protection) unless the contract so provides. But the exact intent of § 105 of the 1976 Act in this regard is not clear. See USCCAN, supra note 15, at 59 (§ 105 deliberately avoids making any sort of outright, unqualified prohibition against copyright in works prepared under Government contract or grant); S & H Computer Sys. v. SAS Institute, 568 F. Supp. 416, 419 (M.D. Tenn. 1983) (government funding of a project does not prohibit a copyright for any developments under the project). See also 1 M. NIMMER, supra note 46, § 5.06[B][2].
A similar economic consideration may apply to the initial publication of research articles by professional journals.\textsuperscript{118} Incentives to undertake expensive and time-consuming research may be increased by allowing scientists to copyright raw data. Furthermore, requiring scientists to replicate experiments and observations in previously explored areas will produce a checking procedure for the original data. Arguably, then, providing copyright protection would contribute to the advancement of knowledge.

However, allowing copyright of research data would more likely impede, rather than encourage, the dissemination of data. First, the replication argument has been discredited.\textsuperscript{119} Second, because "expression" and "fact" merge in research data,\textsuperscript{120} granting copyright protection to the data will have the effect of giving a scientist virtual monopolistic control over facts which would otherwise be part of the public domain. Other scientists would, of course, have the right to reinvestigate a field by conducting their own experiments, but subsequent researchers would not be able to copy data without the first researcher's permission.\textsuperscript{121} If the experiments are too costly and time-consuming to replicate, as is likely with many research topics today, the actual results of any scientific experiment would be monopolized to the detriment of both the scientific and general community.

Furthermore, because little replicative research is done, copyrighting raw data would result in a significant detrimental impact on surveys of scientific subjects. No scientifically useful comprehensive summary of data would be available because any such subject would be a "derivative work"\textsuperscript{122} to which the original copyright owner has exclusive rights.\textsuperscript{123}

\textsuperscript{118} But see Williams & Wilkins Co. v. United States, 487 F.2d 1345 (Cl.Ct. 1973), aff'd by an equally divided Court, 420 U.S. 376 (1975) (per curiam) (extensive photocopying and distribution of articles from medical journals constitutes "fair use" under particular circumstances related to the advancement and dissemination of medical knowledge). Justice Blackmun did not participate, and later suggested that he would disapprove of the "fair use" rationale in this case. See Sony Corp. v. Universal City Studios, 464 U.S. 417, 467 n.16 (1984) (Blackmun, J., dissenting).

\textsuperscript{119} See supra note 61 and accompanying text.

\textsuperscript{120} See supra notes 105-12 and accompanying text.

\textsuperscript{121} "Others are free to copy the original. They are not free to copy the copy." Bleistein v. Donaldson Lithographic Co., 188 U.S. 239, 249 (1903). If scientific research data were copyrightable, there would be no public domain material in the data which other scientists would be entitled to copy.

\textsuperscript{122} "A derivative work is a work based upon one or more preexisting works such as a translation . . . abridgement, condensation or any other form in which a work can be recast, transformed, or adapted." 17 U.S.C. § 101 (1982). The issue of whether derivative works in matters where expression and fact converge should also be unprotectable by copyright law is beyond the scope of this Comment.

\textsuperscript{123} 17 U.S.C. § 106(2).
At best, copyrighting of raw data would lead to repeated, needless experiments to duplicate the copyrighted data. \(^{124}\) Scientists would not be able to extend the research of their predecessors where old research is necessary to establish support for further findings. \(^{125}\) This kind of restrictive control of data would be a great hindrance to the progress of knowledge \(^{126}\) — especially in an enterprise such as scientific research where each scientist relies so extensively upon the contributions of other scientists.

Furthermore, the policy of granting protection as an economic incentive \(^{127}\) is not as central in scientific research as in the realm of directories and other fact-gathering works, because there are incentives and rewards for scientific research other than the commercial exploitation of the collected data. \(^{128}\)

In short, the decision whether to grant copyright protection to research data is informed by conflicting motivations. In general, when there is a conflict between public interest and the potential copyright owner's interest, the former will prevail since the primary objective of copyright protection is to serve the public interest. \(^{129}\) Thus, when "expression" and "fact" merge, \(^{130}\) as in research data, considerations related to the free flow of information should prevail over the other policy.

---

\(^{124}\) Congress expressed concern with the needless repetition of measurements in research and engineering programs in a different setting. See S. REP. NO. 1230, supra note 117, at 2581. Congress' concern was with scientists unknowingly and needlessly duplicating the work of other scientists, not with replications of experiments to check or expand the work of others.


\(^{126}\) Other authors may begin work where the prior author has stopped and may use the copyrighted work as a means of reference to the original sources. See Greenbie v. Noble, 151 F. Supp. 45, 67 (S.D.N.Y. 1957); Sampson & Murdock v. Seaver-Radford Co., 140 F. 539, 540 (1st Cir. 1905). In the words of Professor Chafee: "'The world goes ahead because each of us builds on the work of our predecessors. 'A dwarf standing on the shoulders of a giant can see farther than the giant himself.'" Chafee, Reflections on the Law of Copyright, 45 COLUM. L. REV. 503, 511 (1945).

\(^{127}\) See supra text accompanying notes 117-18.

\(^{128}\) For example, other rewards may include such things as improved reputation within the scientific community and research grants.

\(^{129}\) Shipley & Hay, supra note 17, at 129 n.33. A major consideration in Williams & Wilkins, 487 F.2d 1354, was the possible injury to medical and scientific research of photocopying and distributing individual articles without the copyright owners' permission. See also Twentieth Century Music Corp. v. Aiken, 422 U.S. 151, 156 (1975); United States v. Paramount Pictures, Inc., 334 U.S. 131, 138 (1947). Concerning the economic theory behind copyright protection, see Mazer v. Stein, 347 U.S. at 219; Fox Film Corp. v. Doyal, 286 U.S. 123, 127 (1932); Washington Publishing Co. v. Pearson, 306 U.S. 30, 36, reh'g denied, 306 U.S. 668 (1939).

\(^{130}\) See supra notes 105-12 and accompanying text.
considerations embodied in copyright law.

As the Supreme Court said in *Baker v. Selden*:

The very object of publishing a book on science or the useful arts is to communicate to the world the useful knowledge which it contains. But this object would be frustrated if the knowledge could not be used without incurring the guilt of piracy of the book.\(^{131}\)

**D. Fair Use**

Before leaving the topic of copyright, it should be noted that if scientific research data were given copyright protection, other scientists who copy might not be able to avail themselves of the "fair use" defense against claims of infringement.\(^{132}\) The 1976 Act limits the exclusivity of rights given to a copyright owner for purposes of, *inter alia*, scholarship and research, by permitting "fair use" by others of protected material in certain circumstances.\(^{133}\)

Factors to be considered in determining whether use of a work qualifies as a fair use include:

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.\(^{134}\)

Problems might arise, however, where scientists attempt to use an earlier researcher's data as corroboration of a finding or as the basis for further research. A statistically significant amount of the prior data would have to be utilized to corroborate such a claim. This would probably amount to substantial copying. Since the end result of the second researcher's efforts would probably be a substantially similar collection of data, one would be forced to conclude that there had been infringement.\(^{135}\) Similarly, a useful summary of all the data would be a

---

\(^{131}\) Baker v. Selden, 101 U.S. at 103. Books or research articles would still be copyrightable even if the raw data contained therein were not.

\(^{132}\) Contra D. Nelkin, *supra* note 3, at 43.


\(^{134}\) 17 U.S.C. § 107(1)-(4).

\(^{135}\) See, e.g., Denicola, *supra* note 17, at 531-32. In a compilation case, the Court of Appeals for the Second Circuit stated that "the fruits of another's labor in lieu of independent research obtained through the sweat of a researcher's brow, does not merit copyright protection absent, perhaps, wholesale appropriation." Eckes v. Card Prices Update, 736 F.2d 859, 862 (2d Cir. 1984). "Wholesale appropriation" as the standard for infringement of compilations would be more stringent than the standard for other
protected derivative work, and hence under the control of the original copyright owner.\textsuperscript{136}

Literal application of the factors enumerated in the Act for determining whether a particular use is "fair" discloses even more problems. The first factor, the purpose of the use would favor a claim for fair use since research is listed as a permitted purpose. Yet subsequent scientists are likely to be engaged in the same enterprise, and, in this sense, the scientists are competitors. Although the purpose of the copying is not directly "commercial," it is nonetheless not a totally disinterested, selfless desire to advance knowledge.

With regard to the second factor, the nature of the copyrighted work, one concern will be whether the original copyrighted work is unpublished or not. Under the 1976 Act, publication is no longer the crucial triggering event for copyright protection, but under section 106(3) of the new Act the copyright owner has "the right to control the first public distribution of an authorized copy . . . of his work."\textsuperscript{137} Hence, the fair use defense is not likely to be available if the copyrighted data are unpublished, since use of unpublished data would supplant the copyright owner's valuable right of first publication.\textsuperscript{138}

The third factor, the "amount and substantiality of the portion" of the copyrighted work utilized, will be a major stumbling block to a fair-use defense since any scientifically useful copying will take a significant portion of the data.\textsuperscript{139}

The last factor, the economic effect of the use upon the value and potential market of the copyrighted work, is the most important of the four.\textsuperscript{140} To negate a fair-use defense, "one need only show that if the challenged use 'should become widespread, it would adversely affect the potential market for the copyrighted work.'"\textsuperscript{141} In the case of unpublished data, the effect of unauthorized use of data upon the discovering

\textsuperscript{136}See supra notes 122-23 and accompanying text.
\textsuperscript{137}USCCAN, supra note 15, at 62.
\textsuperscript{138}In Harper & Row, Publishers v. Nation Enterps., the Supreme Court concluded that, while the right of first publication is subject to fair use like other section 106 rights, the unpublished nature of a work is a "key" factor tending to negate a defense of fair use. "Under ordinary circumstances, the author's right to control the first public appearance of his undisseminated expression will outweigh a claim of fair use." 471 U.S. 539, 555 (1985).
\textsuperscript{139}See supra text accompanying note 64.
\textsuperscript{140}See, e.g., Taylor, supra note 17, at 66.
scientist's career may be tremendous since there is no market for duplicative data. In addition, claims to priority and substantiation of work for research grants may be lost.

In short, the fair use defense would probably not be available to scientists who copy another's original research data. Thus, the damaging repercussions of allowing copyright protection for research data would not be abated by the existence of the fair use doctrine.

E. Conclusion

The conclusion from this discussion is that all relevant considerations lead to the same result—there is no justification for granting copyright in research data. Both courts and commentators have suggested that scientific literature in general should receive only limited protection. Clearly, even that limited protection should not extend to the underlying facts and raw research data.

"[W]hen an idea is such that any use of that idea necessarily involves certain forms of expression, one may not copyright those forms of expression, because to do so would be in effect to copyright the underlying idea." Scientific research data fall squarely within this prohibition.

II. MISAPPROPRIATION

Although the policy of free dissemination of information argues against, and ultimately prohibits, copyright protection of raw data, the other policy considerations discussed suggest that at least some measure of protection should be provided for research data. State causes of action might be considered an appropriate alternative for providing such protection. In particular, the states might consider providing a mechanism for protection of unpublished data. Potential claims include conversion, misappropriation, unfair competition, and causes of action based on contract and quasi-contract. This Section considers only those claims which, like copyright, would provide a property interest in the abstract data.

The cause of action most suited to protecting data is a form of unfair competition called "misappropriation"—the taking of the fruit of

142. Sampson & Murdock Co. v. Seaver-Radford Co., 140 F. Supp. 539, 541 (1st Cir. 1905) (dictum) (concerning all arts and sciences); Rubin v. Boston Magazine Co., 645 F.2d 80, 84 (1st Cir. 1981) (limited protection of scientific material when used for scientific, scholarly, news reporting, or like purposes). Cf. 1 M. Nimmer, supra note 46, § 2.11[B], at 2-161 (public interest may justify copying substantial portions of literal form of expression of other kinds of factual works).


144. See Shipley & Hay, supra note 17, at 151.
another's time and effort for competitive advantage. The United States Supreme Court first articulated the doctrine of misappropriation in *International News Service v. Associated Press*.\(^{145}\) In that case, the Court held actionable as misappropriation the conduct of a news-gathering organization in systematically copying and selling to clients fresh foreign news that another service had gathered abroad. The Court further held that the news service which had acquired the news items by its organized expenditure of labor, skill, and money had thereby acquired a quasi-property interest in the news, which would be valid only while the news remained "hot."\(^{146}\) The court observed that no one may claim a monopoly on the gathering or distribution of news which is only the report of information in the public domain,\(^{147}\) but held that even if the competitor acknowledged the source of information, when the competitor endeavors "to reap where [he] has not sown,"\(^{148}\) the "transaction speaks for itself, and a court of equity ought not to hesitate long in characterizing it as unfair competition in business."\(^{149}\)

The *International News Service* approach opens the possibility for the judicial recognition of quasi-property rights in scientific research data. The basis of the Court's recognition of a quasi-property right, however, does not appear to be significantly different from the "sweat of the brow" rationale for copyright in compilation cases.\(^{150}\) Because this rationale has been rejected,\(^{151}\) it should not be used as a basis for providing protection of raw data under the guise of quasi-property rights. In addition, raw scientific data may be sufficiently different from news so as not to justify the recognition of quasi-property rights in that data. Yesterday's news is old hat, so that the quasi-property right expires quickly. It is not nearly so simple to determine whether raw data is "hot" or "cold".

Even if the "sweat of the brow" doctrine were revived, and a suitable test developed for applying quasi-property rights to raw data, the

\(^{145}\) 248 U.S. 215 (1918).
\(^{146}\) *International News Service* held that although:

[n]either party has any remaining property interest as against the public in uncopryrighted news matter after the moment of its publication, it by no means follows that there is no remaining property interest in it as between themselves. \(...\) Regarding the news, therefore, as but the material out of which both parties are seeking to make profits at the same time and in the same field, we hardly can fail to recognize that for this purpose, and as between them, it must be regarded as quasi property, irrespective of the rights of either as against the public.

*Id.* at 236.

\(^{147}\) *Id.* at 241.

\(^{148}\) *Id.* at 239.

\(^{149}\) *Id.* at 240.

\(^{150}\) See supra text accompanying notes 41-47 and 57-63.

\(^{151}\) See supra notes 47 and 49 and accompanying text.
states would not automatically be able to provide such protection to researchers. One must first determine whether federal copyright law has preempted misappropriation doctrine, at least as it might apply to scientific research data.

A. Doctrine of Preemption

Section 301 of the 1976 Act provides that federal copyright law preempts all state law rights and causes of action that are "equivalent" to those protected by the federal Copyright Act. Section 301(b) further provides that federal law does not preempt state causes of action with respect to:

(1) subject matter that does not come within the subject matter of copyright as specified by sections 102 and 103, including works of authorship not fixed in any tangible medium of expression; or

(3) activities violating legal or equitable rights that are not equivalent to any of the exclusive rights within the general scope of copyright as specified by section 106.

Stated affirmatively, the Copyright Act preempts a state action if and only if two conditions are satisfied: (1) the subject matter of the work is within the scope of the Copyright Act; and (2) the protected rights are equivalent to the exclusive rights specified by section 106.

The intention behind section 301 is to preempt and abolish any rights under state law that are equivalent to copyright and that extend to works subject to copyright protection. Whether misappropriation in particular has been preempted by federal law is not clear. Misappropriation was included in the original 1976 copyright bill as an example of a state action not preempted by the new Act. The House amended the bill by deleting the reference to misappropriation but

154. See USCCAN, supra note 15, at 130.
156. See USCCAN, supra note 15, at 24.
157. See H.R. CONG. & ADMIN. NEWS 5810, 5820.
offered no explanation for the deletion.\textsuperscript{158} The Act’s legislative history thus leaves open to dispute whether, and to what extent, section 301 is intended to preempt misappropriation doctrine.\textsuperscript{159} While the Senate adjudged misappropriation to be "nothing more than copyright protection under another name,"\textsuperscript{160} the House determined that misappropriation "is not necessarily synonymous with copyright infringement."\textsuperscript{161} In short, the legislative history is not at all conclusive regarding preemption of state misappropriation law.

Because of the absence of a clear expression of legislative intent on preemption, the following discussion examines the literal wording of section 301 and the two criteria set forth therein to determine whether the Act preempts misappropriation claims against scientists who appropriate for their own use the research data of their colleagues. This Comment concludes that federal copyright law preempts most claims based on the doctrine of misappropriation, thereby leaving scientists without intellectual property protection for their raw research data.

\section*{B. Equivalent Rights}

Section 106 secures to the owner of a copyright the exclusive rights to reproduce the copyrighted work, to prepare derivative works based upon the copyrighted work, to distribute copies of the copyrighted work, and to perform or display the copyrighted work publicly.\textsuperscript{162} Under the usual test a state right is not equivalent to a federal right and is not preempted if it requires proof of some element instead of, or in addition to, those acts enumerated in section 106.\textsuperscript{163}

\textsuperscript{158} See id.

\textsuperscript{159} On the legislative history of section 301, see 1 M. Nim\textsuperscript{m}er, supra note 46, § 1.01[B], at 1-13 to -16. See also Abrams, supra note 155, at 537-550; Fetter, supra note 155, at 34-52.


\textsuperscript{163} See, e.g., Mayer, 601 F. Supp. at 1535 (issue is whether plaintiff’s conversion and misappropriation claims contain an ‘extra element’ which qualitatively distinguishes the actions and their underlying rights from those addressed by copyright law); Rand McNally & Co. v. Fleet Management Sys., 591 F. Supp. 726, 739 (N.D. Ill. 1983), reh’g denied, 624 F. Supp. 604 (1986) (an additional element such as breach of fiduciary duty may take misappropriation claim out of preemption reach of section 301). See also 1 M. Nim\textsuperscript{m}er, supra note 46, § 1.01[B], at 1-11 to -12.

Conversely, state law claims which involve only identical elements are preempted. Orth-O-Vision, Inc. v. Home Box Office, 474 F. Supp. 672, 684 n.12 (S.D.N.Y. 1979) (state law claims preempted by the copyright act unless they involve rights not equivalent to the exclusive rights of section 106); 1 M. Nim\textsuperscript{m}er, supra note 46, § 1.01[B].
This test is surely too permissive, countenancing far more state protection than the dissemination policy behind copyright law would allow. For example, the equivalency test would seem to allow state claims for copying of misappropriated literary works simply because misappropriation claims, unlike copyright claims, all require proof of the plaintiff's considerable time and expense and the defendant's intent to reap a competitive advantage.\textsuperscript{164} Professor Gorman has argued that "[b]ecause it is possible to frame almost any state tort so as to evince a protective policy different from copyright, the proffered analysis would too often interfere with the dissemination policies of the Copyright Act and particularly of section 301."\textsuperscript{165}

Professor Abrams has also suggested that the current approach to determining equivalency is weak, and has proposed that the appropriate way to determine if rights are equivalent is to determine what right is being asserted, rather than to compare its elements of proof.

Whether the antecedent conditions for asserting a right under state law are identical to copyright infringement or diametrically opposed to it is simply irrelevant. The question to ask is whether the right being asserted is one of the exclusive rights listed in § 106. Thus proving that the claimant has invested great time, money, and skill, or that the claimant will suffer harm, is no more germane than proving that the claimant has blue eyes.\textsuperscript{166}

It is clear that the rights sought by scientists through the misappropriation doctrine are exactly the same as those enumerated in the Copyright Act—protection against unauthorized reproduction or distribution of their raw factual material.\textsuperscript{167} Even claims governing the use of unpublished research data would ordinarily involve rights equivalent to the section 106 rights against unauthorized reproduction because a scientist would actually have to publish the misappropriated and previously unpublished data in some form in order to support any conclusion based on the data.\textsuperscript{168}

\textsuperscript{164} Implications for Copyright, supra note 17, at 608. But see Mayer, 601 F. Supp. at 1535-36 (misappropriation claim against copying of snowflake design preempted, notwithstanding plaintiff's showing of her own effort and defendant's immorality); Rand McNally, 591 F. Supp. at 739 (commercial immorality and wrongdoing are not additional elements sufficient to avoid preemption).

\textsuperscript{165} Implications for Copyright, supra note 17, at 608.

\textsuperscript{166} Abrams, supra note 155, at 577.

\textsuperscript{167} See Rand McNally, 591 F. Supp. at 739 (misappropriation claim against copier of highway mileage information not exempt from copyright preemption); Mitchell v. Penton/Indus. Publishing Co., 486 F. Supp. 22, 25-26 (N.D. Ohio 1979) (dismissing misappropriation claim against user of factual information about business records maintenance); cf. Denicola, supra note 17, at 517 n.7 (the appropriation of factual material from non-literary works is nothing more than simply copying).

\textsuperscript{168} See supra notes 135-36 and accompanying text.
At best, section 301 might exempt from copyright protection some claims based on sustained and systematic misappropriation of research data. The legislative history of the Copyright Act directs that:

state law should have the flexibility to afford a remedy (under traditional principles of equity) against a consistent pattern of unauthorized appropriation by a competitor of the facts (i.e., not the literary expression) constituting “hot” news, whether in the mold of *International News Service v. Associated Press*, 248 U.S. 215 (1918), or in the newer form of data updates from scientific, business, or financial data bases.\(^{169}\)

Absent such a pattern of unauthorized appropriation, however, a state action for the misappropriation of scientific research data will reduce to the equivalent of an action against copying or reproduction.\(^{170}\)

Even if the use of research data were not equivalent to any section 106 right, permitting states to protect rights in the use of scientific data clearly goes too far. Because “expression” merges with “fact” in data,\(^{171}\) protecting the use of such data would allow scientists to control material in the public domain which would directly conflict with the fundamental copyright policy of encouraging the free flow of information. This cannot be justified even by a policy providing incentives to undertake scientific research. Allowing scientists to control the use of their research data would inhibit the progress of science by providing individual scientists with too much control over raw facts—even after such facts had been introduced into the public domain.

Thus, with the possible exception of extraordinary cases involving consistent, unauthorized appropriation,\(^{172}\) the rights protected under a state action for misappropriation of scientific data are equivalent to those protected by the 1976 Copyright Act and therefore cannot avoid preemption under the equivalent rights test.

### C. Subject Matter

The "subject matter" test of section 301(b) as applied to research data is more problematic. Section 301(b) permits state protection for subject matter falling outside sections 102 and 103; section 102(b) provides that copyright protection does not “extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or

---


\(^{170}\) *See* 1 M. Nimmer, *supra* note 46, § 1.01[B][2][b], at 1-19 (no substantive distinction between actions for copying and misappropriation).

\(^{171}\) *See* *supra* notes 105-12 and accompanying text.

\(^{172}\) *See* *supra* note 163.
embodied.” Thus, under a literal reading, section 301(b) would permit broad state protection of ideas, processes, etc.173

Moreover, section 301(a) specifies that only “works of authorship that are fixed in a tangible medium of expression and come within the subject matter of copyright as specified by sections 102 and 103” are governed exclusively by the Act. Since facts and data are not “works of authorship,”174 they would not be governed by the Copyright Act, and arguably would be open to state protection.

The difficulty with this literal reading of section 301 derives from the general goal of the Copyright Act. The House Committee Report contains the following passage.

As long as a work fits within one of the general subject matter categories of sections 102 and 103, the bill prevents the States from protecting it even if it fails to achieve Federal statutory copyright because it is too minimal or lacking in originality to qualify, or because it has fallen into the public domain.175

Research data and other bald expressions of facts fixed in a tangible medium of expression seem to fall easily into this preempted category because they are “too minimal” in their expression. In addition, the Supreme Court held in Goldstein v. California176 that only those areas left “unattended” by federal law—areas in which Congress had “drawn no balance”—are not preempted by federal copyright legislation.177 The inclusion of section 102(b) in the 1976 Act strongly suggests that Congress has drawn a balance in the area of facts and intended that the


174. See supra notes 79-112 and accompanying text; 1 M. NIMMER, supra note 46, § 2.11[A] and [E].

175. USCCAN, supra note 15, at 131. See Financial Information v. Moody’s Investor Serv., 808 F.2d 204 (2d Cir. 1986) (fact compilations which were denied copyright protection because of insufficient originality, selection, and arrangement were nonetheless works of authorship under the Act and therefore state misappropriation claims preempted); but see 1 M. NIMMER, supra note 46, § 2.11[E], at 2-167 n.30 (misappropriation of news, i.e., facts, not preempted because news per se is not a “writing” and Congress may not legislate a copyright in facts).

By contrast, the House Committee Report states that data which are not fixed in a tangible medium of expression (e.g., data on the cathode ray tube of a computer terminal) should be afforded state protection. USCCAN, supra note 15, at 132.


free flow of facts not be restrained in any manner. In other words, the area of facts has been covered by congressional action and Congress has deliberately left it unprotected.178 As the Supreme Court stated in Goldstein, "a conflict would develop if a State attempted to protect that which Congress intended to free from restraint or that which Congress had protected."179

With respect to section 301(b), the consequence of the House Committee Report and Goldstein is that fact-expressions fall "within the subject matter of copyright as specified in sections 102 and 103," and thus state protection is preempted. In the words of Professor Gorman:

[w]hen Congress declares in section 102(b) that copyright in such literary work does not "extend to any idea" described, explained or embodied therein, it is not declaring such an idea outside the subject matter of copyright so much as it is affirmatively declaring—as clearly as it can, and for the clearest reasons—that ideas are free to be copied, adapted and disseminated, and that no court is to construe the federal copyright monopoly as inhibiting that freedom. The implication for state law is equally clear: neither can the states. . . . Far from leaving facts, ideas and the like "unattended," to borrow a term from the Goldstein case, Congress has very much attended to them in section 102(b), and has declared them to be free as the air.180

Thus, facts have been deliberately excluded from copyright protection and therefore the states cannot protect them.181 Permitting state protection of simple factual expressions would create "vague borderline areas between State and Federal protection."

---

178. The House Committee Report states: (section 102(b) in no way enlarges or contracts the scope of copyright under the present law. Its purpose is to restate, in the context of the new single Federal system of copyright, that the basic dichotomy between expression and idea remains unchanged. USCCAN, supra note 15, at 57. As Professor Abrams has said concerning this passage, "[t]here is certainly nothing in this comment that can conceivably be taken as endorsing such a drastic change as to allow the states to redefine the public domain to restrict access to fact." Abrams, supra note 155, at 564.

179. 412 U.S. at 559.

180. Implications for Copyright, supra note 17, at 604-05. Gorman’s comments echo Brandeis’ dissent in International News Service v. Associated Press: "the general rule of law is, that the noblest of human productions—knowledge, truths ascertained, conceptions, and ideas—become, after voluntary communication to others, free as the air to common use." 248 U.S. 215, 250 (1918) (Brandeis, J. dissenting). See also Shipley & Hay, supra note 23, at 165; Abrams, supra note 155, at 541-42.

181. Where historical facts, themes, and research have been deliberately exempted from the scope of copyright protection to vindicate the overriding goal of encouraging contributions to recorded knowledge, the states are pre-empted from removing such material from the public domain. Hoehling v. Universal City Studios, 618 F.2d at 980 (citing Sears, Roebuck & Co. v. Stiffel Co., 376 U.S. 225 (1964) and Compco. Corp. v. Day-Brite Lighting, 376 U.S. 234 (1964)).
contrary to the intent of section 301.182 The creation of these "vague borderline areas" would also be contrary to the general congressional intent to provide a "single Federal system" of statutory copyright protection which "would greatly improve the operation of the copyright law and would be much more effective in carrying out the basic constitutional aims of uniformity and the promotion of writing and scholarship."183

Thus, it appears doubtful that states could provide a property interest protection for scientific research data, because such protection fails both prongs of the preemption test: first, the rights of importance to scientists in protecting raw data are equivalent to those provided by section 106 of the Act; second, raw scientific data fall within the subject matter considered by the Act. Therefore, there is no room for non-federal property interest in scientific research data.

CONCLUSION

A scientist conducting an experiment and gathering data is not an author of an original work in a sense relevant to copyright. It may seem anomalous that no protection is available under the copyright laws despite the ingenuity and labor expended in creating and carrying out scientific experiments, but from the point of view of copyright law, the scientific researcher is simply gathering the work of another author: nature. The 1976 Copyright Act was not designed to protect scientific data. Furthermore, the cases and secondary authorities on fact-gathering works do not indicate that the Act should be adapted to protect such data. In addition, because scientific data fit squarely into the category of expressions Congress intended to leave unprotected by copyright, the states may not extend protection to such data under the doctrine of misappropriation.

In sum, this Comment has argued that a property interest would not be the proper vehicle for protecting intangible scientific research data. However, this does not mean scientists are necessarily without recourse in asserting rights over their research results. The House Committee Report gives "invasion of personal rights" as an example of a cause of action not equivalent to copyright.184 Conversion, trespass,
misrepresentation, and breaches of contract and of trust are other examples included in the Report.\textsuperscript{185} Other potential actions not mentioned in the Report include unfair competition and false designation of the origin of a work under the Lanham Act.\textsuperscript{186} A District Court has even held that an action concerning trade secrets is not preempted when the material was not copyrighted.\textsuperscript{187} Perhaps these causes of action concerning tangible and intangible property will help secure the rights of scientists in their raw data and protect the integrity of the research process without unduly restricting the free flow of scientific information.

\textsuperscript{185} USCCAN, supra note 15, at 132. The House Committee Report also states: "[n]othing contained in section 301 precludes the owner of a material embodiment of a copy . . . from enforcing a claim of conversion against one who takes possession of the copy . . . without consent." \textit{Id.} at 133. \textit{See also} United States Trotting Ass'n v. Chicago Downs Ass'n, 665 F.2d 781 (7th Cir. 1981) (recovery for the physical taking of a particular fact-document not preempted by the 1976 Copyright Act).

\textsuperscript{186} 15 U.S.C. § 1125(a) (1982) (trademark law). This section creates a federal tort for unfair competition based on "false designation of origin or other false representation used in connection with the sale of a product." Metric & Multistandard Components Co. v. Metric's Inc., 635 F.2d 710, 713 (8th Cir. 1980).