
Celia Delano Moore

Follow this and additional works at: http://repository.jmls.edu/jitpl

Part of the Computer Law Commons, Intellectual Property Law Commons, Internet Law Commons, Legal Writing and Research Commons, Privacy Law Commons, and the Science and Technology Law Commons

Recommended Citation

http://repository.jmls.edu/jitpl/vol4/iss2/3

This Article is brought to you for free and open access by The John Marshall Institutional Repository. It has been accepted for inclusion in The John Marshall Journal of Information Technology & Privacy Law by an authorized administrator of The John Marshall Institutional Repository.
OWNERSHIP OF ACCESS INFORMATION: EXPLORING THE APPLICATION OF COPYRIGHT LAW TO LIBRARY CATALOG RECORDS

by CELIA DELANO MOORE*

TABLE OF CONTENTS

I. INTRODUCTION ............................................ 306
II. TRADITIONAL CATALOGING AND COPYRIGHT LAW. 307
   A. Copyright In the Catalog Record ..................... 307
      1. Nature of the Catalog Record ...................... 307
         a. Descriptive Cataloging .......................... 308
         b. Subject Cataloging ................................ 310
      2. Subject Matter of Copyright ........................ 311
         a. Fixation ......................................... 312
         b. Authorship ....................................... 312
         c. Originality ...................................... 312
      3. The Catalog Record As the Subject of Copyright ... 314
   B. Copyright In the Library Catalog ..................... 315
      1. Nature of the Library Catalog ...................... 315
      2. Copyright In Compilations .......................... 316
         a. Statutory Provisions ............................. 316
         b. Arrangement ...................................... 317
         c. Aggregation ....................................... 317
         d. Selection .......................................... 318
      3. The Catalog As a Compilation ....................... 319
III. IMPACT OF DEVELOPMENTS IN COMPUTER TECHNOLOGY .............................................. 321
   A. Changes In Cataloging ................................ 321
      1. Development of Bibliographic Utilities ............ 321


305
I. INTRODUCTION

When a library adds a work to its collection it adds a record to its catalog, describing the physical characteristics of the work and its content in accordance with standard conventions. Traditionally, libraries have sought to minimize duplication of this descriptive effort by sharing catalog records. Recent technological developments, such as the growth of on-line cataloging using large data bases of catalog records input by many libraries, have made catalog sharing...
Ownership of Access Information

Routine. New technology has also made possible new uses of catalog records by libraries and by commercial entities. Libraries have recently become sensitive to the questions these developments raise about ownership of bibliographic data.

This Article explores questions regarding the application of copyright law to library catalog records, questions applicable to other data bases as well. Copyright law itself has been adjusting to changed technological conditions; it provides no clear or easy answers to the problem.

To provide the necessary background, this Article first describes the traditional nature of the catalog record and the catalog itself and describes how copyright law applies to those forms. It then outlines the effect of technological developments, particularly the computer, on cataloging activity and copyright law.

The next section explores problems in applying present law to library bibliographic data by tracing the ownership rights the law provides in a situation involving complicated (but typical) uses of library bibliographic data. These difficulties have been avoided to a large extent by private negotiation, as is pointed out in a subsequent discussion of typical contractual arrangements.

The doctrine of fair use provides flexibility in the application of copyright law that is particularly necessary where libraries are involved. The doctrine also provides a satisfactory balance between the rights of the creator of the bibliographic data and the public need for access to knowledge.

II. TRADITIONAL CATALOGING AND COPYRIGHT LAW

A. Copyright in the Catalog Record

1. Nature of the Catalog Record

In order to determine whether a catalog entry falls within the category of works protectible by copyright, one must examine the typical catalog entry and the process by which it is created. The individual entry is shaped largely by its intended use as part of a comprehensive catalog of the library's collection. The purposes of the familiar library dictionary catalog were well stated by Charles Cutter in 1904:

1) To enable a person to find a book of which is known either a) the author b) the title c) the subject.
2) To show what the library has d) by a given author e) on a given subject f) in a given kind of literature.
3) To assist in the choice of a book g) as to the edition (biblio-
For the catalog to achieve these objectives the book must be described in a way that differentiates it from all others. Librarians define two separate parts of this process of identification: descriptive cataloging and subject cataloging.

a. Descriptive Cataloging

"Descriptive cataloging is that phase . . . which is concerned with the identification and description of a book and the recording of this information in the form of a cataloging entry . . . . without reference to its classification into a subject area or to the assignment of subject headings . . . ." Its purposes are to "1. present the title and certain other information commonly used for the citation of the work 2. describe the physical object 3. provide information relating to the work's bibliographic history 4. identify the nature, scope, artistic form, intended audience, and contents of the work . . . ."

On first view, the process of describing a book seems extremely simple, more or less one of transcription. In fact, however, descriptive cataloging in the United States has been an area of intellectual concern and debate for almost a century, for which a number of different sets of rules have been formulated and revised. Those rules currently governing the practice of most libraries, the second edition of the Anglo-American Cataloging Rules, implemented in 1981, occupy over 600 pages, and were the subject of intense professional controversy before their final adoption.

One area in which disagreement is possible is in the selection of information to be presented in the description. Author and title are obviously important, but what about identification of the illustrator, or the number of blank leaves at the end of the book, or the fact that the book is part of the publisher's "Books for Millions" series? The desire to use the catalog entry as a surrogate for the book itself, requiring an exhaustively detailed description of its physical charac-


2. B. WYNAR, supra note 1, at 29.


characteristics, must be weighed against the use of the entry as a finding tool, a pointer to the book in question.

Of course, each library can choose which elements it wishes to describe, regardless of the practice at other libraries. Militating against such an individualistic approach is the desire for predictability, benefitting the mobile patron and librarian, and "uniformity which allows universal application."\textsuperscript{5} Uniform descriptive rules allow one library's catalog record of a book to be used by all others with little modification, since the same elements will be found regardless of the identity of the cataloging library. Under current practice, the following descriptive elements are included: title and statement of responsibility (author, editor, translator), edition, imprint (place, publisher, date), physical description (pagination, presence of illustrations, height in centimeters), and identity as part of a series.

Once the descriptive elements have been selected, they must then be described. The concepts of "author" and "title" require detailed and complex exegeses in order for catalogers to reach the same results in describing the constitution of Costa Rica, a collection of speeches given at a conference, a corporation's annual report, or a German translation of the Koran.\textsuperscript{6}

Another problem involves "entry," or the selection of access points. For example, will the record be found in only one place in the catalog, or will duplicate records (or shortened forms thereof) be placed under other headings in the catalog? Here the record will very much reflect the nature of the catalog of which it is destined to become a part. In the past, because book catalogs had space limitations and cards had to be copied manually, there was generally but one occurrence of the record, the "main entry."\textsuperscript{7} Abbreviated "added entries" were sometimes made. With the rise of the printed, easily duplicated card, it was theoretically possible to make as many entries as desired simply by typing a heading atop each copy of the card. But exigencies of space and manageability limit the number of entry points feasible in any manual system. Thus, choosing the headings that will be used becomes significant; again the desire for uniformity has led to governing rules whose application requires judgment.\textsuperscript{8}

\textsuperscript{5} B. WYNAR, supra note 1, at 29.
\textsuperscript{7} P. DUNKIN, supra note 1, at 45.
\textsuperscript{8} See ANGLO-AMERICAN CATALOGING RULES, supra note 4, at 277, for a sampling of these governing rules.
b. **Subject Cataloging**

Whatever the complexities of descriptive cataloging may be, they have always been considered to be less challenging than those of subject cataloging, which includes classification into a subject area (choosing a call number) and assignment of subject headings.

Library classification "arranges the expression of knowledge as preserved in written records . . . with the specific purpose of providing an adequate subject approach to the existing collection." Broadly speaking, the librarian uses classification to place books on a given subject together, preferably near books on similar or related subjects. A number of different schemes of library classification exist; the two most commonly used in the United States are the Dewey Decimal and the Library of Congress classifications. Each has become a highly complex system with an ever-proliferating array of subdivisions. The Library of Congress classification schedules occupy twenty-eight separate volumes.

When assigning a book to the one place in the classification where it most properly belongs and is most likely to be sought, the librarian has little more guidance than such general exhortations as "classify a book where it will be most useful," "in the most specific subject that will contain it," "according to the intent of the author." Browsers can attest to the important consequences of the exercise of the cataloger's discretion.

Classification provides a systematic approach to subject matter. Subject headings "give a topical approach, collecting all aspects of a subject under its . . . name." Many libraries choose their headings from two standard lists, each of which is several hundred pages long. Choosing subject headings from similar terms is said to be governed by the principle that one selects the "most specific entry." This is not much of a limit on discretion.

A catalog entry, then, describes the form and content of a work in accordance with a body of standards. From the brief discussion given of the work of the cataloger it should be apparent that each catalog entry results from the exercise of considerable skill and judgment acquired through education and experience. As such, it

---

9. See, e.g., P. Dunkin, supra note 1, at 69.
10. B. Wynar, supra note 1, at 200.
11. See, e.g., id. at 220-21.
12. Id. at 263.
13. The standard lists are the Library of Congress Subject Headings (9th ed. 1980), and the Sears List of Subject Headings (10th ed. 1972).
15. Cataloging of this sort is, in most libraries, performed by professional librari-
fits fairly comfortably within the boundaries of permissible subjects of copyright protection under existing law.

2. Subject Matter of Copyright

The United States Constitution gives Congress the power "to promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." Congress enacted the first copyright law in 1790, and after major revisions in 1831 and 1870, the Copyright Act of 1909 was enacted. Problems soon occurred in applying the statute to the numerous technological advances in reproduction and communication of images and sounds. These problems led to repeated proposals for revision. Thus, after numerous studies and much debate, Congress enacted the Copyright Act of 1976, which took effect January 1, 1978. This Act does away with the previous dual system under which state common law copyright protected unpublished works, while federal statutory protection attached upon publication and compliance with notice requirements. Under the new Act there is a single unified system of federal statutory copyright protection beginning at the moment a work is created.

Section 102 of the new Act defines the subject matter protected by the statute:

(a) Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device. Works of authorship include the following categories: (1) literary works...

\[\]
a. Fixation

Section 101 explains that "[a] work is ‘fixed’ in a tangible medium of expression when its embodiment in a copy or phonorecord, by or under the authority of the author, is sufficiently permanent or stable to permit it to be perceived, reproduced or otherwise communicated for a period of more than transitory duration." Because of this broad definition, the fixation requirement is seldom difficult to meet.

b. Authorship

The fixed work must be a "work of authorship." Congress purposely left that phrase undefined, although it provides a rather full list of works included as works of authorship. The first category, "literary works," is defined in section 101 as including works, other than audiovisual works, expressed in words, numbers, or other verbal or numerical symbols or indicia, regardless of the nature of the material objects, such as books, periodicals, manuscripts, phonorecords, film, tapes, disks, or cards in which they are embodied. The House report on the Act noted: "The term 'literary works' does not connote any criterion of literary merit or qualitative value: it includes catalogs, directories, and similar factual, reference, or instructional works and compilations of data."

c. Originality

A significant portion of the litigation concerning the subject matter of copyright involves the definition of "originality." This term was purposely left undefined, as Congress intended "to incorporate without change the standard of originality established by the courts under the present copyright statute." The 1909 Act did not define originality, nor did it expressly require it. But "courts uniformly inferred the requirement from the fact that copyright protection may only be claimed by 'authors' . . . . It was reasoned that since an author is 'the beginner . . . or first mover of anything . . . ,' it follows that a work is not the product of an author unless the work is origi-
nal." The House report does note that the standard of originality "does not include requirements of novelty, ingenuity, or esthetic merit." Thus a work need not be unique, though it must be independently created. "[If] by some magic a man who had never known it were to compose anew Keats' Ode On a Grecian Urn, he would be an 'author,' and, if he copyrighted it, others might not copy that poem, though they might of course copy Keats." A very low level of originality is necessary. Indeed, the most commonly cited formulation of the requirement states: "All that is needed to satisfy both the Constitution and the statute is that the 'author' contributed something more than a 'merely trivial' variation, something recognizably 'his own.' Originality in this context 'means little more than a prohibition of actual copying.' However, "the contribution by the author must amount to more than a trivial variation." "Trivial" may refer to the quality or to the quantity of the original contribution. Thus, the Copyright Office regulations state that words and short phrases are not subject to copyright, no matter how creative they may be. Furthermore, longer works are not protectable if they consist "entirely of information that is common property, containing no original authorship."

Copyright protection is not extended to ideas, but only to their expression. Section 102(b) provides: "In no case does copyright protection for an original work of authorship 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 embodied in such work." Facts in themselves are not copyrightable since "they may be discoverable, but are not created by an act of authorship." However, the literal form of expression of the facts cannot be copied, except where that expression must

30. M.M. Business Forms Corp. v. Uarco, Inc., 472 F.2d 1137, 1139 (6th Cir. 1973) (holding a business form not copyrightable where its provisions were merely trivial variations of standard legal phrases).
33. Miller v. Universal City Studios, 650 F.2d 1365 (5th Cir. 1981). For a discussion of this case, see M. Nimmer, supra note 26, at § 2.11[A].
34. Rubin v. Boston Magazine Co., 645 F.2d 80, 83 (1st Cir. 1981) (holding plain-
necessarily be formulated in a given manner.\textsuperscript{35} The distinction between expression and idea is simply stated but has created great difficulty in application.

3. \textit{The Catalog Record As the Subject of Copyright}

Catalog records quite readily meet the statutory criteria for copyrightable subject matter. They are tangibly fixed in one of the media within the statute, whether it be book, card, or machine-readable disk. They fall easily within the statutory definition of literary works, since they are "expressed in words, numbers or other verbal or numerical symbols," and need not meet any requirement of literary quality.\textsuperscript{36} A slightly more difficult issue is that of originality. The physical description of the book is essentially a listing of facts which arguably must necessarily be formed in a given manner. However, the mode of expressing, for example, authorship of a work, could have a number of variants. Furthermore, there is an even stronger argument for original creativity in the selection of the facts to be included in the description.\textsuperscript{37} But it is not necessary to examine the extent to which selection of facts may manifest the original creativity necessary for copyright, since a catalog record consists not only of the physical description of the work but also of a classification number and headings representing its subject matter. In a sense, the subject of the book is a fact which is discoverable, not created. However, this fact can be expressed in a great variety of ways. Even from a finite list of possible subject headings and call numbers, there is not one manner in which the description must be formulated. Thus the cataloger's description of the contents is protectable

\textsuperscript{35} Consumers Union of the United States, Inc. v. Hobar Mfg. Co., 199 F. Supp. 860, 861 (S.D.N.Y. 1961) (denying protection to material copied by defendant's sales bulletin where the expressions "describe facts and nothing more and could hardly have been stated in any different fashion"); accord, PIC Design Corp. v. Sterling Precision Corp., 231 F. Supp. 106, 111 (S.D.N.Y. 1964) (finding no infringement in trade catalog entries where "to some extent the substantial similarities which do appear as to the arrangement . . . are the natural result of the subject matter. There are a very few ways of listing such information and the general method used by both plaintiff and defendants is logical and to be expected.").

\textsuperscript{36} 17 U.S.C. § 101 (1982).

\textsuperscript{37} One could counter that the selection has already been once done by the formulation of cataloging rules and that the creator of the individual record who follows them need exercise little or no selective judgment as to the elements to be included. The question of selection of facts as originality is discussed infra text accompanying notes 55-59.
though the idea expressed is not.\textsuperscript{38}

Library catalog records have apparently never faced a litigated test of originality for copyright purposes, but at least one case has held individual descriptive entries in a trade catalog to be copyrightable. In that case, the plaintiff "collected data on the supplies and conceived a lucid and forceful description by arrangement of the essential data . . . [t]he originality consists in the description of each item."\textsuperscript{39} Whether "lucid and forceful" or not, library catalog records do contain more than a trivial variation on extant factual information. It appears quite certain that a catalog record is copyrightable as an original work of authorship fixed in the necessary tangible form.\textsuperscript{40}

B. COPYRIGHT IN THE LIBRARY CATALOG

Unlike the majority of copyrighted works, an individual catalog record has not been intended to stand alone. It is shaped by and produced for the catalog of which it will become a part. A library may have minimal investment in the individual entry, but the catalog itself is an indispensable part of the library's operation.\textsuperscript{41}

1. Nature of the Library Catalog

Two views of the catalog's purpose are possible. The earliest and still the most significant for many users is that of an inventory and finding list. "Early catalogs listed the items one by one as they were stored . . . with no reference to any arrangement but storage."\textsuperscript{42} In addition to organizing books as physical entities, the catalog records may reflect their nature as embodiments of an author's work. Under this second view the library catalog is "an organizing instrument, which collects works under the agent responsible for their genesis and assembles bibliographic units into literary

\textsuperscript{38} See Rubin v. Boston Magazine Co., 645 F.2d 80 (1st Cir. 1981). For a discussion of the body of cases supporting this accepted proposition see M. Nimmer, \textit{supra} note 26, at § 2.11[B].

\textsuperscript{39} Markham v. A.E. Borden Co., 206 F.2d 199, 202 (1st Cir. 1953).

\textsuperscript{40} If a work is produced by an employee as a work for hire, the employer is considered the author for purposes of the statute, and will be the owner of the copyright, unless otherwise agreed. 17 U.S.C. § 201(b) (1982). A "work made for hire" is defined as "a work prepared by an employee within the scope of his or her employment." 17 U.S.C. § 101 (1982). In the typical library situation there is little doubt that cataloging would be work done within the scope of the librarian's employment. Thus the library would own the copyright in the cataloging records its librarians created.


\textsuperscript{42} P. Dunkin, \textit{supra} note 1, at 5.
units." Expressed as early as 1674, this idea found an ardent and influential proponent in Anthony Pannizi at the British Museum in the mid-nineteenth century, and is still reflected in the Cutter objectives for the catalog which were noted above. Under this view "the catalog is not simply an index to the collection. In a sense, by its ability to create the collocation of entries, it assists in defining the collection. By an examination of its various sequences the catalog permits us to determine what has been written about a particular subject, how a particular idea evolved, how a particular author's thinking evolved, how a particular author's work affected others, what manifestations of an author's work the library owns, etc."

The view one takes of the catalog will affect one's analysis of the application of copyright law to it. It is undisputed that copyright protection may be extended to compilations; but the rationale for this extension, and consequently the scope of the protection, is unclear.

2. Copyright In Compilations

   Section 103(a) states that "the subject matter of copyright as specified by section 102 includes compilations and derivative works." A compilation is defined as "a work formed by the collection and assembling of preexisting materials or of data that are selected, coordinated, or arranged in such a way that the resulting work as a whole constitutes an original work of authorship." The House report notes that a compilation results regardless of whether the individual items compiled have been or could have been subject to copyright. Thus, while individual facts may not be copyrighted, statutory protection extends to compilations of facts, even if the facts themselves are in the public domain. The statute reflects prior

43. S. MALINCONICO & P. FASANA, supra note 41, at 3.
45. Malinconico, The Library Catalog In a Computerized Environment, in Freedman & Malinconico, supra note 44, at 49.
47. H.R. REP., supra note 23, at 57; 1976 U.S. CODE CONG. & AD. NEWS at 5670. A subset of compilations are collective works, defined as those "in which a number of contributions, constituting separate and independent works in themselves, are assembled into a collective whole." 17 U.S.C. § 101 (1982). A catalog could be deemed a collective work, as an assembly of separate copyrighted catalog records, but it is a compilation whether or not the records are copyrighted.
b. Arrangement

The common rationale for protecting the compilation of facts but not the underlying facts themselves is that the compiler's arrangement of the facts is protectable expression. As commentators have pointed out, however, this rationale fails to provide protection in the case where a person simply rearranges a compiler's data—a process which may require minimal effort. Not surprisingly, a number of courts have responded to this inequity by finding the labor of compilation to be protectable in itself.

c. Aggregation

A line of cases suggesting that laborious compilation merits protection was originated in and is still best exemplified by the Ninth Circuit's holding in *Leon v. Pacific Telephone & Telegraph.* There the defendant used data from the plaintiff's alphabetical telephone directory, rearranging it by telephone number. The court found that the defendant had infringed upon the plaintiff's copyright, referring in its analysis to the labor and expense undertaken by plaintiff in publishing its directory and deeming those efforts worthy of copyright. This economic rationale for copyright in compilations has been criticized: "To accord copyright protection on this basis alone distorts basic copyright principles in that it creates a monopoly in public domain materials without the necessary justification of protecting and encouraging the creation of 'writings' by 'authors.'"

---

48. A seminal quotation is from Learned Hand: "The right to copyright a book upon which one has expended labor does not depend upon whether the materials which he has collected consist or not of matters which are *publici juris,* or whether such materials show literary skill or originality, either in thought or in language, or anything more than industrious collection." *Jeweler's Circular Publishing Co. v. Keystone Publishing Co.*, 281 F. 83, 88 (2d Cir. 1922) (affirming the validity of a copyright in a catalog of trademarks), *cert. denied,* 259 U.S. 581 (1922). See M. *Nimmer,* supra note 26, at § 3.02, for additional cases.

49. *See,* e.g., Schroeder v. William Morrow & Co., 566 F.2d 3 (7th Cir. 1977) (holding defendant's directory, which copied the gardening catalog's "selection, ordering and arrangement" of names and addresses, an infringement).


51. 91 F.2d 494 (9th Cir. 1937).

52. *Id.* at 485-86.

53. M. *Nimmer,* supra note 26, at § 3.04. Under Nimmer's view, plaintiff could bring an action for unfair competition against the compiler but should not be protected by copyright law. A discussion of alternatives to copyright for protection of intellectual property is unfortunately outside the scope of this Article. *See infra* note 184.
Nevertheless, courts continue to cite Leon as the authority for holding that a subsequent compiler cannot copy data "to save himself the labor and expense incurred by the prior compiler."\(^{54}\)

**d. Selection**

At least one commentator has advocated explicit recognition of an economic basis for copyright protection in compilations.\(^ {55}\) Most courts and commentators feel more comfortable phrasing this view in terms of authorship, since the statute is so phrased. For example, one commentator suggests that a "promising approach is to find authorship in the act of aggregating isolated pieces of information. The particular collection of data would thus itself be a work of authorship."\(^ {56}\) However, even courts following the Leon principle of protecting laborious compilations have hesitated to find authorship in the mere act of aggregation. Instead, they focus on the work of collection as a selection rather than mere aggregation of facts, and have generally found that selection of data is sufficiently creative to merit protection. Thus, in an early case, the defendant was held to have infringed upon the plaintiff's social register to the extent that it copied plaintiff's selection of the families to be included rather than making its own choices.\(^ {57}\) Recently, the Dow Jones lists of key stocks were held copyrightable "due to the effort and judgment exercised in their composition... [The lists]... evidence a high degree of selectivity and subjective judgment."\(^ {58}\)

It is quite clear then that a compiler's particular selection and arrangement of data may be protected. In some jurisdictions the data itself may be held protected by copyright from a defendant's use which takes unfair economic advantage of plaintiff's labor of compilation.\(^ {59}\)

---


\(^{55}\) Squires, Copyright and Compilations in the Computer Era: Old Wine in New Bottles, in TECH. AND COPYRIGHT 205 (G. Bush & R. Dreyfuss eds., 2d ed. 1979). The economic considerations involved in copyright protection of data base compilations will be discussed infra § IV.

\(^{56}\) Denicola, supra note 50, at 530.

\(^{57}\) List Publishing Co. v. Keller, 30 F. 772, 773 (S.D.N.Y. 1887).


\(^{59}\) See infra notes 219-21 and accompanying text.

It should be noted that the great bulk of the compilation cases are not in fact concerned with the copyrightability of plaintiff's work, but rather with whether defend-
3. The Catalog As a Compilation

How do these principles apply to the library catalog? The statute provides that "the 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." What is the material contributed by the author of the library catalog?

Apparently no case involving a library catalog has been litigated. The closest "catalog" case is Hotten v. Arthur, a nineteenth century British case in which an injunction was issued barring the defendant from copying portions of the plaintiff bookseller's catalog of rare books. Rather summarily, the Vice Chancellor stated that the catalog was protected:

The catalogues written by plaintiff were not mere dry lists, but contained descriptions written by the plaintiff of the books offered for sale. Supposing these descriptions had been composed by some author and printed in a bookseller's catalogue, it was clear that the author would have a copyright in them, and no other book seller could reprint such descriptions without infringing that right. Mr. Hotten was both author and publisher of these catalogues, but that could make no difference in his rights; he had expended considerable time and labour on this work, and when once the compilation was printed, no one had a right to copy it.

The language of Hotten points up problems in analyzing a library's catalog in relation to copyright. The court states that the catalog is copyrightable because it is not merely a dry list but contains descriptions of the books. But if one focuses on the catalog as a collection of individual copyrighted descriptive entries, then the element of originality contributed by the catalog as a whole seems to ant in fact copied it as opposed to doing independent compilation. They thus tend to focus on questions such as how substantially similar the works are, or what use defendant actually made of plaintiff's work in compiling his own. Though the issue of copyrightability has been the central focus of only a few case holdings, courts have then routinely cited them as settled propositions. Case law is comparatively clear but it does not rest upon a great body of law. See Gorman, Copyright Protection for the Collection and Representation of Facts, 76 Harv. L. Rev. 1569 (1963); Squires, supra note 55, at 214.

60. 17 U.S.C. § 103(b) (1982).

61. Hotten v. Arthur, 32 L.J.Ch. (n.s.) 771, 772-73 (1863). The case has been cited only once since its decision, for the proposition that a trade catalog may be copyrighted. See Grace v. Newman, 19 L.R.-Eq. 623 (1875). The case seems to dispose of the preliminary objection that no originality is present in a library catalog because its contents are constrained by the library's collection. Though the bookseller's catalog was necessarily descriptive only of the books in his stock the description of that limited set of books evidently was supposed to provide ample room for original selection and arrangement of facts relating to them.
vanish from the dictionary catalog as a work of compilation, for the idea of alphabetical arrangement of entries is not itself copyrightable. The common image of the "catalog" as essentially a piece of furniture into which cards are filed leads one to see it as quite different in character from the compilations involved in List and Dow—no compiler stands by the drawer selecting which cards to include. Even the labor argument is less persuasive for a compilation done a little at a time over a span of many years. Thus, under the view of the catalog as a finding list, there is significant doubt as to whether it is a proper subject for copyright.

A better view may be the one which explicitly takes account of the very long time frame involved, seeing the catalog as one ongoing compilation. Selection and arrangement of data to be included in its component parts is continually occurring as catalogers follow the rules of entry and standard headings established for the whole catalog. Under this view of the catalog as an organizing instrument, original material is continually "contributed" to the existing entries of the compilation, making it a "work of authorship" under section 103. The catalog rules of entry and heading provide the structure into which the entries fall. Thus while each catalog entry reflects the general rules, the catalog as a whole embodies their creative expression. The catalog under such an analysis does meet the criteria necessary for copyright. Arrangement and selection are occurring and certainly the requisite labor is present.

All analyses aside, one can also rely pragmatically on the large body of precedent recognizing copyright in other types of catalogs.

62. See Oberman, Copyright Protection for Computer Produced Directories, 22 COPYRIGHT L. SYMP. (ASCAP) 1 (1977); Denicola, supra note 50, at 528; Gorman, supra note 59, at 49. Discussion generally centers around Leon, where the court held that an alphabetically arranged phone book for the San Francisco area was infringed by a numerically organized directory for the same area. Oberman notes that one could view the plaintiff's "idea" as a collection of names and his "expression" as alphabetical arrangement. The court however did not focus on the arrangement as protectable expression, but rather the labor involved as meriting protection, implying that in fact alphabetical arrangement per se was not a protectible original arrangement. 63. From this point of view the creativity involved in producing a single entry seems to be less, for the cataloger is bound by the rules of the catalog entity as a whole. Still, because of the "skill and judgment" involved in the selection of subject terms and classification, if not in the bibliographic description, the individual catalog record is copyrightable.

64. See, e.g., Schroeder v. William Morrow & Co., 566 F.2d 3 (7th Cir. 1977). For a history of the developments leading to recognition of copyright in trade catalogs, see Jackson, Compilations As Subjects for Copyright, 31 KY. L.J. 231 (1943). See also M. Nimmer, supra note 26, at § 3.02.
It seems quite clear therefore that the library's catalog is a copyrightable compilation.

III. IMPACT OF DEVELOPMENTS IN COMPUTER TECHNOLOGY

A. CHANGES IN CATALOGING

So far "catalog" and "catalog record" have been spoken of in abstract terms, assuming a basic irreducible entity independent of the medium in which it is embodied. In the last fifteen years, however, new forms have developed for library catalogs and catalog records which may change the basic concept of each (if they have not already). The computer has had a tremendous impact on library cataloging; the question of ownership of bibliographic data has become of significant concern to libraries largely because of this impact. To explain this new concern, a description of developments in cataloging technology is necessary.

1. Development of Bibliographic Utilities

Librarians have always understood that for each of them to catalog the same work involves a very costly duplication of effort, and have generally felt that bibliographic data should be shared. As early as 1853, Charles Jewett of the Smithsonian Institution library proposed that one stereo-type plate be made for each title cataloged by that library, so that other libraries could easily obtain copies of its cataloging. The American Library Association proposed centralized distribution, and in 1901 the Library of Congress began to make copies of its printed cards available for purchase by any library. In the same year, the Library of Congress also began administrating the National Union Catalog. Over 800 libraries have submitted records to the catalog regarding titles which they have cataloged. From 1856 on, the National Union Catalog has appeared

---

in book form.\textsuperscript{68}

The Library of Congress thus became established as the major producer of bibliographic data long before the use of computers in library applications. In the 1960's, the Library developed the Machine Readable Cataloging (MARC) format, "a data structure and conventions for the representation of bibliographic data."\textsuperscript{69} The actual content of a MARC record consists of two types of fields:\textsuperscript{70} control fields which contain coded information about the item described (e.g., frequency of a serial, or language); and variable length fields containing text such as the title or publisher. Each field has an identifying tag. Variable fields also have indicators, which identify the type of content within them. Cataloging for 200,000 of the titles the Library of Congress cataloged each year thus became available in machine-readable form. Nevertheless, even the Library of Congress does not catalog all materials published, and not all of its cataloging was produced in machine-readable form. Thus, libraries still faced, as they had before MARC, the task of individually cataloging books which many other libraries had previously cataloged.

In 1971, the Ohio College Library Center (OCLC), a non-profit organization serving the Ohio College Association, introduced the first version of an on-line cataloging system. MARC tapes were made available on-line to participating libraries; for titles not found in MARC, a participating library entered its own original cataloging. Any other libraries later searching for cataloging of that work would find a ready-made record.\textsuperscript{71} The high cost of original cataloging could thus be significantly reduced.\textsuperscript{72} OCLC enjoyed a phenomenal success. Regional library networks initially thought of replicating the system, but instead became brokers for their member libraries.
in joining the OCLC system. By July 1982, OCLC had over 3000 library users and a data base of over eight million records. In addition to its original function as an on-line cataloging system with the capacity of producing printed cards to each cataloging library’s specifications, OCLC has added interlibrary loan (ILL) and serials check-in components. OCLC was the first of the “bibliographic utilities.”

Original cataloging was significantly decreased for two-thirds of the respondents in the Markuson survey. Markuson, supra, at 88. The Hewitt respondents reported that original cataloging had become necessary for almost no titles in smaller libraries (budgets of $100,000 or less) and for less than 20% of the titles in larger libraries. Hewitt, supra, at 272.

Although original cataloging costs decreased, libraries did not necessarily save on the overall cataloging process. Georgia State University reported that its cataloging costs rose slightly after joining OCLC, from $275,000 to $284,000. Landram, Cataloging: OCLC Terminal Plus Printer, 21 LIBR. RESOURCES & TECH. SERVS. 147, 154 (1977). Ohio State University, in contrast, reported that “[i]f inflation were not a factor, there would have been a substantial decrease in unit costs.” Morita & Gapen, A Cost Analysis of the Ohio College Library Center On-Line Shared Cataloging System in the Ohio State University Libraries, 21 LIBR. RESOURCES & TECH. SERVS. 286, 301 (1977). The Markuson survey found that “[s]ome libraries claimed savings, others claimed increases, and still others reported little or no change,” and concluded that “no clear cost trends due to OCLC use were discerned.” Markuson, supra, at 31-32. More recently, the general view seems to be that use of a bibliographic utility does not lead to overall cost reductions. See DeGennaro, Libraries and Networks in Transition, 106 LIBR. J. 1045, 1047 (1981).

Libraries overwhelmingly felt that bibliographic utilities were worth any increase in cost due to the benefits of greatly reduced cataloging time as well as secondary benefits such as increased information about other libraries’ holdings. Markuson, supra, at 32, 36; Hewitt, supra, at 273.

OCLC contracts with the network, which in turn contracts with individual member libraries. As one commentator describes it: “The functions of regional networks are contracting, marketing, education, evaluation, surveys, accounting, maintenance, and research and development.” Evans, Constituency Concerns in OCLC Management: User, Library, Network, OCLC, in REQUIEM FOR THE CARD CATALOG: MGMT. ISSUES IN AUTOMATED CATALOGING 141, 142 (D. Gore, J. Kimbrough & P. Spyers-Duran eds. 1979).

Bibliographic Utilities, supra note 71. See also Whither OCLC?, AM. SOC’Y FOR INFO. SCI. BULL., Aug. 1982, at 16.

The basic concept of an online bibliographic utility is that each participating library has access to both source bibliographic records, for example, the records made available through the Library of Congress MARC distribution service, and the original cataloging contributed by other participating libraries. Access to the data base is generally through a CRT terminal connected via dedicated telephone lines to the utility’s computer system. When needed, a search is made of the data base and the CRT displays any records that match. The library then may make some modifications to the record, such as adding local call numbers and holdings information. This information is then stored in the computer and, depending upon the needs of the library, cards or
A number of large libraries have developed their own in-house computer systems for cataloging or processing. The BALLOTS system developed at Stanford became the Research Libraries Information Network (RLIN) when Stanford joined the Research Libraries Group (RLG), a non-profit organization of large research libraries. Similar to OCLC in its function as a bibliographic utility, RLIN currently has 382 library users, but is governed by the 29 members of RLG. These large research libraries perform a significant percentage of original cataloging. There are two other active bibliographic utilities in North America.

The growth of the bibliographic utilities has lent impetus to the librarians' dream of a national network. Ideally, a library, linked via a terminal to a computer and the computer in turn linked either to other computers or to a communications network, would have access to all of the machine-readable data in America. Thus, every library would have access to shared cataloging information, holdings information for ILL, and be better able to respond to patron reference questions.

Libraries have taken steps toward that end. The Library of Congress has established a Network Development Office to participate in projects and studies on networking. A Network Advisory Committee, with representatives from the regional library networks, gives policy advice, while technical assistance is provided by a Network Technical Architecture Group. The Council on Library Resources used foundation grant money to start a Bibliographic Services Development Program to try to encourage complementary efforts among bibliographic utilities, networks, large private and national libraries and vendors, and funded a study on linking of OCLC and RLIN.

other products may be provided by the utility to the library. The result of such a system is that each library has less original cataloging to do. Also, whatever original cataloging is done is shared with other libraries.


76. Bibliographic Utilities, supra note 71, at 627. The article contains a useful description of RLIN's present nature.

77. These are the University of Toronto Library Automation Systems (UTLAS) and Washington Library Network Computer Service (WLN). Bibliographic Utilities, supra note 71, at 627.

78. Matthews, supra note 75, at 757. See also Avram, supra note 65, at 360.

However, despite these developments, the proliferating literature about library networking recognizes the technological, organizational, and economic issues which must be resolved before a truly national network, which even then is more likely to be a mosaic of decentralized systems than a monolithic entity, can be developed.80

2. New Uses of Catalog Records Made Possible

As yet, only sixteen percent of all libraries use a bibliographic utility, though most of the larger academic libraries are among the users.81 Even in those libraries, the intellectual process involved in original cataloging has remained essentially unchanged. The same rules govern bibliographic description, and traditional subject heading lists and call numbers are used. The cataloger does have to fit each into an appropriate field of the computer record format with proper identifying tags, and there are a few fields for information that were not traditionally included in catalog records, such as the ISBN. Even so, the content has not changed.

In one respect cataloging has changed. For the majority of libraries that still maintain a card catalog while using a bibliographic utility, cards representing the cataloging record will arrive from the utility and become part of the "compilation" that is the library's catalog. In these libraries, the record also becomes part of an invisible library catalog consisting of all the records that the library has placed in the bibliographic utility and all the records of other libraries to which it has attached its holding symbol. The utility regularly produces a tape for the library embodying that "invisible" catalog.82 Some libraries use these tapes to create their own computer output microform or on-line catalogs.

The catalog record is not only contained in that library's catalog,
regardless of what form that catalog takes. Instead, the record also becomes part of the utility's data base. Many other libraries will use the record, modifying it as they wish and making it part of their catalog, just as the originating library adopted other libraries' records for its own catalog. In a sense, what is happening is nothing new, because for years libraries have adapted Library of Congress catalog records for their own use. However, Library of Congress records are not subject to copyright, while the original records of individual libraries are. Now that it is possible for records to be used by other institutions, libraries have become aware of ownership issues related to cataloging records.

In March 1980, the Network Advisory Committee held a meeting devoted to the question of ownership and distribution of bibliographic data. Participants submitted suggestions of pertinent issues in advance of the meeting. A partial list of them rings with the notes of uncertainty.

Who should own MARC level records created with federal monies that become part of the data bases run by networks, service utilities, or institutions which use state funds? The Library of Congress, though the intellectual owner of records, considers MARC records in the public domain: should distribution by sale make these records the property of the purchaser? If an institution inputs a record into a utility, it intellectually owns the record; however, if the institution does not pay for the record, should this be interpreted to mean that since the utility did not charge for use of the system, it (the utility) has paid for the record? Should the record belong to the utility as well as the inputting institution? If a member library uses a record that is already in the data base and pays for the record, should the record belong to the second library as well as to the utility? What restrictions can be placed on the use of data by the owner? Should a library have sole authority to release its data for use in another network? Should a fee be charged to reimburse an institution or agency for input costs, particularly if a large unique collection is involved? If a library gives away data it created, should it then have to pay for its use in another network?

Part of the uncertainty is due to the potential uses of catalog records contained in a data base. Computer technology has made possible repeated use of a record as part of a different data collection in each use. To provide a focus for the issue, three examples of proposed (and actual) uses of records will be described.

The Harvard Law Library, a special member of RLG, inputs original catalog records into the RLIN data base, which also contains

---

OWNERSHIP OF ACCESS INFORMATION

original records from other RLIN member libraries as well as MARC records from the Library of Congress. Each library receives tapes from RLIN containing the records that the library has cataloged on RLIN either by providing its own records as input or by modifying cataloging input from other libraries. A compiler obtains these tapes from several RLIN libraries and produces a list of their recent acquisitions (eliminating all duplicate records), which he offers for sale to any interested buyer on a continuing basis. Each library whose tape was used is given statistics on acquisitions generated from its tape, but must pay if it wishes to obtain either the combined acquisition list or its own acquisition list.

The second example concerns a compiler who expands his operation. Tapes from the Government Printing Office (which catalogs its publications using another bibliographic utility, OCLC) are added to the RLIN library tapes, along with tapes from Information Access Corporation (IAC), a commercial service which indexes legal periodicals and newspapers. This additional data is compatible with the RLIN tapes since both catalogers and the indexer put the data into the MARC format and use Library of Congress subject headings. The compiler may also add the results of searches in data bases available through systems such as Lockheed's Dialog. The resulting data base, compiled from many sources, will be offered by the compiler as a resource in which he will do searches for a fee.  

The third example is an application of data base searches obtained for a fee. Under such a fee arrangement, the Harvard Law Library receives a print-out from the compiler each month containing the results of searches of his data base on requested subjects. It distributes the results to the university's faculty members as a current awareness service. It also prints a copy of the search results for review by students in the library reading room. Reaction has been favorable, so the library makes copies available to students for a fee covering the printing and copying cost, and then to the public at a price designed to let the library make a profit.

The question of who has rights in the bibliographic data involved in these situations is not hypothetical. Each party wants to have as much access to useful information as possible, but each wants to avoid infringing upon another's copyright, if not for ethical

84. Alternatively, the compiler might offer direct access to his data base. If a library using its own in-house computer system searched his and other data bases, and stored, merged, and published the results as a bibliography, problems of ownership rights analogous to those presented by situations (2) and (3) in the text would arise. Although the discussion will focus only upon situations (1) through (3), its implications for similar situations should be kept in mind.
then for practical reasons such as eliminating the costs of uncertainty and litigation.

Questions of copyright law in such situations have not been litigated. In large part this may result from the nature of the institutions facing these questions, and the avoidance of potential problems by privately negotiated agreements. But in part the lack of helpful case law reflects the general confusion about copyright law as related to computers, which even the most recent attempt at statutory clarification has not entirely eliminated.

B. Changes In Copyright Law

1. Development of Law Relating to Computer Applications

Just before the Copyright Act of 1909 was passed, the Supreme Court was faced with the problem of applying existing copyright law to new technology: a player piano roll. The Court held that copying of a song by means of the perforated roll did not infringe the copyright in the musical composition because the music could not be visually perceived from the piano roll. "It is not susceptible of being copied until it has been put in a form which others can see and read." The concept of visually perceptible copies was adopted by the 1909 Copyright Act.

During the early years of computer development, therefore, it seemed that copyright protection would not be available for a work recorded in electronic impulses.

In 1964, however, the Copyright Office announced that it would accept computer programs for registration; it required deposit of the programs "in language intelligible to human beings" as a precondition to registration. The announcement defined computer programs as "either a set of operating instructions for a computer or a compilation of reference information to be drawn upon by the computer in solving problems." The second clause might be interpreted as covering machine-readable data bases. But data bases are not really intended to be drawn upon by the computer in solving problems. "Rather, they consist of information available in machine-readable form which is sought out in a process directed by a set of operating instructions—a program—and presented as the solution to the problem." In any event, complying with the deposit requirement by producing print-outs of large and continuously

87. Announcement SMI-47 from the Office of the Register of Copyrights, COPYRIGHT OFFICE CIRCULAR 31D (1965) [hereinafter cited as Announcement].
88. Id.
89. Squires, supra note 55, at 220.
changing data bases would not be practical.90

Besides the practical difficulty of copyrighting a data base, the statutory authority for copyright was, as the Register of Copyrights recognized, "dubious."91 Early attempts at revising the Copyright Act to reflect computer technology ran into controversy. In 1964, a Preliminary Draft for Revised United States Copyright Law contained a provision giving the copyright owner the exclusive right to reproduce the work "in any form in the programming or operation of an information storage and retrieval system."92 Later versions eliminated that language.93 The Register of Copyrights explained that it was better to state the general concepts of copyright ownership rights in broad terms that would allow for adjustment to future developments.94 In fact, the provision had generated continuing controversy. Academic users and information scientists felt that imposition of copyright liability for input or use of works in systems would stifle research and development and use of computer systems; authors and publishers wanted protection against unauthorized data entry that might result in electronic images taking the place of printed copies. Users lobbied for either statutory exemption of some computer uses, or explicit recognition that some uses would be fair use.95 Both sides made proposals for establishing voluntary licensing systems.

90. Id.
91. Announcement, supra note 87, at 56. The announcement recognized that the questions of whether a program was a copyrightable "writing" of an author and whether reproduction of the program in the form used by the computer was a regular "copy" were "doubtful questions" under the statute. But it stated that registration would be "in accordance with its policy of resolving doubtful issues in favor of registration."

95. The positions are well summarized in a 1967 Judiciary Committee report:
It has been argued on behalf of those interested in fostering computer uses that the copyright owner is not damaged by input alone, and that the development [sic] of computer technology calls for unrestricted unavailability of unlimited quantities of copyrighted material for introduction into information systems... these interests recommended at least a partial exemption for input. On the other side, the copyright owners stressed that computers have the potential, and in some cases the present, capacity to destroy the entire market of authors and publishers. They consider it indispensable that input, beyond fair use, require the consent of the copyright owner.
At further hearings, witnesses expressed growing recognition that the issue (together with issues related to photocopying) might cause tremendous delay in passage of the new Copyright Act. Finally, in 1974, legislation creating the National Commission on New Technological Uses of Copyrighted Works (CONTU) was enacted. The Commission was to study and make recommendations on "the reproduction and use of copyrighted works of authorship: (A) in conjunction with automatic systems capable of storing, processing, retrieving, and transferring information, and (B) by various forms of machine reproduction, . . . and (2) the creation of new works by the application or intervention of such automatic systems of machine reproduction."  

In the meantime, section 117 of the proposed Copyright Act was added to freeze the existing state of the law.  

Notwithstanding the provisions of sections 106 through 116 and 118, this title does not afford to the owner of copyright in a work any greater or lesser rights with respect to the user of the work in conjunction with automatic systems capable of storing, processing, retrieving, or transferring information, or in conjunction with any similar device, machine, or process, than those afforded to works under the law, whether title 17 or the common law or statutes of a State, in effect on December 31, 1977, as held applicable and construed by a court in an action brought under this title.  

Thus, despite passage of the new Act in 1976 (taking effect January 1, 1978), a large area of law relating to computers remained unclear. However, certain problems were eliminated by the new Act. The concept of fixation expressed in sections 102(a) and 101 eliminated the old White-Smith requirement that a copy be visually perceptible; unauthorized reproduction in any form would be considered an infringing copy. Furthermore, section 102 was intended to afford copyright protection to computer programs and data bases as literary works "to the extent that they incorporate authorship in the programmer's expression of original ideas, as distinguished from the ideas themselves."  

A machine-readable data base compilation

---

96. Act of Dec. 31, 1974, Pub. L. No. 93-573, § 201(b), 88 Stat. 1873. Twelve commissioners were appointed in 1975, representing authors and publishers, copyright users, and the general public. The Commissioners split into four committees dealing respectively with photocopying, computer software, computer data bases, and computer-created works.  

97. Copyright Act of 1976, Pub. L. No. 94-553, 90 Stat. 2565 (repealed 1980). Repeal of this provision as a result of CONTU's work is discussed infra text accompanying notes 102-04.  

could be copyrighted if it met the traditional standards for originality required for hard copy compilations.99

The remaining question was whether a copyright owner had the exclusive right to reproduce (or license reproduction of) his work in machine-readable form, and when this right must be recognized. Viewed in the context of compilations and data bases, the question under section 117 of whether A could input B's copyrighted catalog into A's data base regardless of B's copyright, and the question of what A's proprietary rights in his own data base were against one using it, were to be governed by the law existing at the time of the Copyright Act enactment. Thus, the White-Smith definition of "copy" would still apply and copying would have been an infringement if in eye-readable form but not if in machine-readable form. The problems this partial freeze created were especially acute for computer programs.100 When CONTU issued its final report in 1978 it recommended a new provision that would clearly protect computer programs, but would allow the possessor of a copy of a program to make another copy under certain circumstances.101

The Commission concluded that data bases should be given copyright protection "equivalent to the protection accorded compilations in traditional hard copy format."102 Once section 117 was repealed, however, no new legislation would be necessary to clarify that status.

The new Copyright Act, in the absence of the limited moratorium imposed by section 117, deals effectively with questions related to copyright protection for automated data bases. For example, under the provisions of section 106, the copying or input of a data base or any other work of authorship embodied in a computer-readable medium is an exclusive right of the copyright owner.103

In 1980, Congress followed CONTU's recommendation and repealed section 117, replacing it with the suggested provision.104

99. Squires, supra note 55, at 221.
101. Final Report, supra note 95, at 12.
102. Id. at 39.
103. Id. at 38.
104. 17 U.S.C. § 117 (1982) states that:
Notwithstanding the provisions of section 106, it is not an infringement for the owner of a copy of a computer program to make or authorize the making of another copy or adaptation of that computer program provided: (1) that such a new copy or adaptation is created as an essential step in the utilization of the computer program in conjunction with a machine and that it is used in no other manner, or (2) that such new copy or adaptation is for archival purposes only and that all archival copies are destroyed in the event that continued possession of the computer program should cease to be rightful.

The copyright status of machine-readable data bases can now be stated with some certainty. Data bases are included in the definition of literary works which may be subjects of copyright.\(^{105}\) There is a notice requirement, however. "Whenever a work protected under this title is published in the United States or elsewhere by authority of the copyright owner, a notice of copyright as provided by this section shall be placed on all publicly distributed copies from which the work can be visually perceived, either directly or with the aid of a machine or device."\(^{106}\) A work is "published" when copies are distributed "to the public by sale or other transfer of ownership, or by rental, lease or lending. The offering to distribute copies or phonorecords to a group of persons for purposes of further distribution, public performance, or public display, constitutes publication. A public performance or display of a work does not of itself constitute publication."\(^{107}\) CONTU explained:

a data base proprietor, by display alone, could make the data base available to users, without having published the data base. The same would be true where the proprietor leased a tape containing the data base directly to a user and placed that user under explicit restrictions prohibiting disclosure or transfer. Under these circumstances, the failure to place copyright notice on the data base ... would jeopardize no rights the proprietor might have. If, however, the proprietor authorized transferees to distribute copies or make available displays of the data base, publication would be accomplished and the notice and registration requirements of the law would take effect.\(^{108}\)

Any exact copies prepared in accordance with the provisions of this section may be leased, sold, or otherwise transferred, along with the copy from which such copies were prepared, only as part of the lease, sale, or other transfer of all rights in the program. Adaptations so prepared may be transferred only with the authorization of the copyright owner. The legislative history of the amendment, concerned with impact on computer programs, contains no significant discussion of copyright with respect to data bases in connection with repeal. It seems reasonable to assume that CONTU's analysis of the effect of repeal can be accepted.


108. Final Report, supra note 95, at 43. One CONTU attorney suggested that display on a terminal might constitute publication if the terminal user could also print from the terminal. "Thus the prudent proprietor of data bases might well wish to include the notice on all displays as well as on all tangible copies of his data base." Bigelow, Copyright in Computerized Data Bases, Computer L. Serv. (Callaghan) § 4.3, art. 3, at 8 (1977) (quoting Christopher Meyer, an attorney with CONTU, at the Computer Law Association Meeting in Mar. 1977).
Omission of notice from a publicly distributed copy will not necessarily invalidate the copyright. Under Copyright Office regulations notice can be affixed for "works reproduced in machine-readable copies ... from which the work cannot ordinarily be visually perceived except with the aid of a machine or device" by having it appear on each "visually perceptible print-out," or on the user's terminal at sign-on, or continuously displayed on the terminal, or even in visible form as a label on whatever container the machine-readable copy is housed in (e.g., a cartridge or reel).

Copyright registration is not mandatory and is not a condition of copyright protection. It is, however, a condition to certain remedies for copyright infringement and a prerequisite to the commencement of a copyright infringement action. Registration can be obtained by delivering a deposit copy of the work, application, and fee to the Copyright Office. Since the registration of a work is not mandatory, the deposit of a copy with the Register is also not mandatory.

In order to protect himself as fully as possible, a data base copyright owner would want to register the work and arrange for notice of copyright to appear on either the terminal with each use of the data base or any print-out produced from use of the data base. Under section 408, a deposit copy must accompany a registration application. The Register cannot exempt materials entirely from this requirement, but can limit the deposit to appropriate identifying material, which has been done for data bases. The deposit now required is "the first and last 25 pages or equivalent units of the work if reproduced on paper ... or, in the case of automated data

109. 17 U.S.C. § 405(a) (1982) provides that omission will not invalidate the copyright if it is from a relatively small number of copies, if the work has been registered within five years of the publication without notice and a reasonable effort has been made to correct the omission after discovery thereof, or if the omission was in violation of an express written requirement that proper notice was a condition of the copyright owner's authorization for the public distribution of copies of the work. However, an innocent infringer who relied on the absence of a notice may be free from liability under 17 U.S.C. § 405(b) (1982). The 1976 provision is in sharp contrast to prior law where forfeitures could result from "unintentional or relatively unimportant omissions or errors in the copyright notice." H.R. Rep., supra note 23, at 143; 1976 U.S. Code Cong. & Ad. News at 5759.

110. 37 C.F.R. § 201.20(g) (1983).

111. Id.


113. Section 407 does make mandatory the deposit for the use of the Library of Congress of two copies of any work published with notice of copyright in the United States. However, the section gives the Register of Copyrights authority to exempt categories of material from this deposit requirement, and the Register has used this power to exempt machine-readable works. 37 C.F.R. § 202.19(c)(5) (1983).
bases comprising separate and distinct data files, representative portions of each separate data file consisting of either 50 complete data records from each file or the entire file, whichever is less.\textsuperscript{114} In the latter case, a descriptive statement must accompany the deposit containing information such as "the subject matter involved, the origin(s) of the data, and the approximate number of individual records within the file" as well as "a description of the exact contents of any . . . copyright notice employed in the work and the manner and frequency with which it is displayed."\textsuperscript{115}

It is not clear how long this one-time identification will be sufficient for registration of a dynamic data base whose contents are constantly changing. CONTU addressed fears that by constantly updating the data base a proprietor could claim copyright of the work in perpetuity regardless of the statutory term of seventy-five years applicable to the data base under the new statute, speculating that "the proprietor of a data base would have to register for copyright each update of the work, just as the proprietor of a telephone directory obtains copyright in new editions of the work."\textsuperscript{116}

One author notes that nothing in the new Act "speaks directly to the unique problems confronted by those who desire to register claims to copyright in materials such as data bases which are constantly being modified to include new material."\textsuperscript{117} His suggestion that "the Register . . . rule that registration for a data base would protect all modifications in that data base for a specified period of time at the end of which prior registration . . . would have to be updated" has not been adopted.\textsuperscript{118} Thus, this area of the law remains unclear.

The new statute expressly grants the owner of the data base exclusive rights to do or to authorize someone "(1) to reproduce the copyrighted work in copies or phonorecords; (2) to prepare derivative works based upon the copyrighted work; (3) to distribute copies . . . of the copyrighted work to the public . . . ; (4) . . . to perform the copyrighted work publicly; and (5) . . . to display the copyrighted work publicly."\textsuperscript{119}

\textsuperscript{115} Id.
\textsuperscript{116} \textit{Final Report}, \textit{supra} note 95, at 41-42.
\textsuperscript{117} Squires, \textit{supra} note 55, at 228.
\textsuperscript{118} Id. at 229. In fact, when the current regulations were issued for comment, no comments on the subject were received. 43 Fed. Reg. 41,979 (1978). Current practice manuals do not discuss the issue as a problem.
\textsuperscript{119} These exclusive rights are "subject to sections 107-18." Sections 110-16 and 118, however, deal with uses of types of material other than literary works and are
IV. PRESENT COPYRIGHT LAW AS IT APPLIES TO CATALOG RECORDS IN DATA BASES

A. RESULTS UNDER CURRENT PROVISIONS GOVERNING COPYRIGHT OWNERSHIP RIGHTS

1. Rights of Owners of Underlying Data and Compilers of Data

Given regulations for notice and deposit which recognize the special problems of the machine-readable form, data bases fit quite comfortably into current provisions governing compilations in general. Discussions of the rights of data base proprietors generally assume that the matter contained in the data base is not itself composed of copyrighted components, but is rather like a directory composed of individual unprotectable facts. What constitutes an infringement of the proprietor's rights is therefore generally discussed in terms of the law that has arisen from directory and trade catalog cases. The law as it applies to catalog records in data bases can be summarized as follows.

a. Rights of Compiler Where Underlying Data Is Not Copyrighted

Where catalog records composing a data base are not copyrighted, the underlying records are not protected. However, the compiler's contribution, which consists of selection, arrangement, and possibly labor of aggregation, is protected from copying. As noted above, it is difficult to find "selection" in a compilation of catalog records unless one views the catalog as a unified entity reflecting rules governing selection of descriptive elements to be included. The same is true of a data base, though selection can be found under the same reasoning if one assumes that all contributing libraries follow the same cataloging standards. In fact, despite the standardized MARC formats employed, the literature about OCLC at least suggests that disparate content standards are a problem. If standards were completely disparate, the bibliographic utilities could well be found not to be products of "selection," especially because unlike many other data bases they are not the input of a single

120. See, e.g., Keplinger, Copyright and Information Technology, 15 ANN. REV. INFO. SCI. & TECH. 3 (1980); CRC Systems, Inc., Impact of Information Technology on Copyright Law in the Use of Computerized Scientific and Tech. Information Systems, in TECH. AND COPYRIGHT, supra note 55, at 137, 177-78.

121. For a discussion predating the 1976 Copyright Act, see Oberman, supra note 61.

122. See, e.g., Bibliographic Utilities, supra note 71, at 627. If standards were completely disparate, the bibliographic utilities could well be found not to be products of "selection," especially because unlike many other data bases they are not the input of a single
sourcing above, but it is more difficult to find in a data base since the records are not actually located in the computer's memory or storage in any sequence meaningful to a human being. They can, however, be retrieved as if filed in a given alphabetical order. Thus, one can rather easily argue that the effective arrangement necessary is present.

So far as the labor of collection is concerned, the utility's activities in developing and maintaining the system clearly make it an even stronger candidate for protection than the industrious directory compilers in the *Leon* line of cases.123

The problems of finding selection and arrangement in a bibliographic data base are present in any other data base where data has been contributed by more than one person; it is significant that Congress nevertheless expressly stated the intention to make copyright protection possible for data bases that are compilations.124 Given that fact and the additional factor of labor expended, which justified copyright under the case law, the data bases of bibliographic utilities are almost certainly copyrightable as compilations.125

Nonetheless, the component parts of a compilation are not protected by the compilation's copyright. Thus, where the underlying records in a data base are not copyrighted they are not protected by the compilation's copyright, and anyone may copy them individually as long as the bibliographic utility's compilation copyright is not infringed.126 Catalog records, however, are copyrightable. The data base of the bibliographic utility may thus belong to that subclass of compilations known as a "collective work," i.e., one that consists of contributions which themselves constitute works capable of being copyrighted.127 Before examining copyright in the machine-readable catalog situation, therefore, one more aspect of copyright law must be explored: the respective rights of the owner of the collective work and of the owner of an underlying work included in the

---

123. Squires argues that in fact the economic rationale is strongest for protecting data bases. Squires, *supra* note 55, at 230.
125. OCLC has copyrighted its data base as of Dec. 27, 1982. See infra note 147.
126. Infringement by copying can be shown where defendant has access to the copied work and there is "substantial similarity" between the two. Application of this standard presents "without doubt the most vexed question in the law of copyright," and tests are "of necessity vague." N. BOORSTYN, *supra* note 112, at § 10:14. In compilation cases the question revolves around whether the compiler's idea (facts, content) has been copied permissibly or if his expression (arrangement, selection) has been copied impermissibly. See M. NIMMER, *supra* note 26, at § 8.
127. See M. NIMMER, *supra* note 26, at § 3.02.
collection.128

b. Rights of Compiler and Owner of Underlying Data Where Underlying Data Is Copyrighted

Section 103(b) makes it clear that two different copyrights are involved:

The copyright in a compilation or derivative work 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 an exclusive right in the preexisting material. The copyright in such a work is independent of, and does not affect or enlarge the scope, duration, ownership, or subsistence of, any copyright protection in the preexisting material.129

Section 103(a) further notes that “protection for a work employing preexisting material in which copyright subsists does not extend to any part of the work in which such material has been used unlawfully.”130 In working out the permutations of any factual situation it is important to bear in mind the independence of the copyrights.131

The producer of a new work (“N”) copied from a compilation may copy only component underlying works (“U”), or only the compiler’s original contribution to the compilation (“C”), or both. Briefly sum-

128. In the discussion of copyright of traditional catalogs as compilations, this problem was not dealt with because the library owner of the individual records would also be the owner of the compiled catalog. With the advent of bibliographic data bases, however, the creator of the individual record will not be the owner of the compiled data base. Thus, it becomes important to ascertain the respective rights of each owner.


130. 17 U.S.C. § 103(a) (1982). Boorstyn notes that the provision does not necessarily require permission from the owner of the underlying work; use of his work even without his consent might be “lawful” if in accordance with a statutory exemption. N. BOORSTYN, supra note 112, at § 2.1.

131. One must also realize that under the new Copyright Act each of the five exclusive rights given to the copyright owner may be transferred and owned separately. Section 201(d)(2) provides for this divisibility of copyright (a departure from past law) and states that the owner of any particular exclusive right is entitled, to the extent of the right, to all the protection and remedies accorded to the copyright owner by the Act. 17 U.S.C. § 201(d) (2) (1982). Thus, if a compiler is given the exclusive right to reproduce a particular underlying work, he is the owner and is the person who may sue if the right to reproduce is infringed. N. BOORSTYN, supra note 112, at § 10.7. Under § 501(b) “the court may require such owner to serve written notice of the action . . . upon any person shown . . . to have or claim an interest in the copyright . . . [or] any person whose interest is likely to be affected by a decision in the case,” and may require joinder of such a person. Thus, the compiler might be required to join the copyright owner of the underlying work if that owner would be affected by the decision. 17 U.S.C. § 501(b) (1982).
marized, the law is as follows.\textsuperscript{132}

\textit{N uses only U's}. Where the U's are copyrighted, the owner of a U can sue N for copyright infringement even though U is copied from C. The separability of the two copyrights is shown in \textit{Russell v. Price}, a recent Ninth Circuit case.\textsuperscript{133} In \textit{Russell}, the owners of a copyright in a play brought an action for infringement against persons who were distributing copies of a motion picture. The owners of the copyright in the play had granted third parties the right to use the play to produce a motion picture. Although the copyright of the motion picture had expired, the copyright owners of the play were held to have a cause of action against the distributors for infringement of those portions of the motion picture taken from the play. Thus, using our works and actors, if N used any of C, then the owners of the U's comprising that part of C would have an action against N for infringement, regardless of the copyright status of C.

If the owner of a U has made the owner of a C the exclusive licensee of one or more of his ownership rights in U, the owner of the C, as the owner of those particular rights, has standing to sue N for infringement.\textsuperscript{134} The owner of U may be joined as a party, but does not have to be.

\textit{N uses only C}. The owner of C has a cause of action against N. Where an author who compiled quotations sued for infringement of the compilation, the court held him to be the proper party, regardless of who owned the copyright in the underlying quotations, because he owned the copyright in the compilation.\textsuperscript{135} It seems unlikely that N would take only material constituting the compiler's original contribution (e.g., a foreword and editor's notes to a collection of essays), so this case may seldom arise. Note that the compiler's arrangement (as opposed to selection) of U's is not a protectable subject of copyright under section 102(b).

\textit{N uses material comprised of a U and a C}. Where U is copyrighted, the owners of U and C have causes of action against N for infringement of their respective rights in U and C, as noted in (1) and (2) above. The owner of U cannot sue for infringement of C other than that part which is his work. Similarly, the owner of C

\textsuperscript{132} For a summary of the relevant precedent, see M. Nimmer, \textit{supra} note 26, at §§ 3.05, 3.07.

\textsuperscript{133} 612 F.2d 1123 (9th Cir. 1979), \textit{cert. denied}, 446 U.S. 952 (1980).

\textsuperscript{134} Swarovski America Ltd. v. Silver Deer Ltd., 537 F. Supp. 1201 (D. Colo. 1982). In \textit{Swarovski}, the plaintiff was granted an exclusive right to manufacture the copyrighted figurines, and the court stated that it thus had standing to sue for infringement of the right of manufacture. \textit{Id.} at 1205. It notes that joinder of a person having or claiming an interest in the copyright is permissive. \textit{Id.} at 1206.

cannot sue N for infringement of U in itself unless he is assigned rights in U or is made an exclusive licensee of U's owner. 136

So far, only the rights of the owners of U and C as against a third party have been discussed. What about their rights as against each other? Generally, these rights will be governed by a contract embodying the terms on which the owner of U consents to the use of his work in C. If the owner's consent to the use of U is limited in either time or medium, then the compiler must abide by those limits. Otherwise, the use will be an infringement. 137 If U enters the public domain, then the restrictions on its use will lapse. 138

2. Results In Library Applications

How would these general principles apply in the three representative situations described above of use of library catalog records in the bibliographic utilities? It will be instructive to examine briefly results produced by the statute (unmodified by contractual arrangements) in a situation where all the parties exercise their rights of ownership to the utmost, copyrighting every protectible work. As will be shown, the results are unworkable, if not nonsensical.

In the first of the three hypothetical situations, the Harvard Law Library inputs original catalog records into the RLIN data base, which also contains original records from other RLIN member li-

136. Eden Toys, Inc. v. Florelee Undergarment Co., 526 F. Supp. 1187 (S.D.N.Y. 1981) (a non-exclusive licensee does not have standing to sue for infringement of the right, since he is outside the language of § 501(d), rev'd in part, 697 F.2d 27 (2d Cir. 1982)). Thus, if the owner of a U grants permission to more than one compiler to reproduce his work, none of the compilers can sue N for infringement of the underlying work.


138. It has been shown that if the compilation enters the public domain the copyright in U is unaffected, though there may be unlimited use of the original portions of the compilation. Nimmer explores at length a line of cases which seem to hold that once a derivative work is created pursuant to a valid license to use the underlying material, a new property right springs into existence with respect to the entire derivative work, so that even if the license is thereafter terminated the proprietor of the derivative work may nevertheless continue to use the material from the underlying work as contained in the derivative work. Id. at § 3.07[A]. One recent case seems to espouse this view. See Rohauer v. Killiam Shows, Inc., 551 F.2d 484 (2d Cir. 1977). Nimmer criticizes Rohauer's reasoning at length.

Rohauer was decided before the new Copyright Act became effective, and the language of the Act concerning the separability of copyrights in derivative and collective works from underlying works seems clearly opposed to its result. It therefore seems defensible to disregard it for purposes of the brief summary of applicable law given here.
libraries as well as MARC records from the Library of Congress. The Harvard Law Library is the author of the original catalog record its employee cataloger creates. To copyright it and obtain maximum protection Harvard or any other inputting library would attach a notice of its copyright to the record and register the record with the Copyright Office by depositing a copy of it. Even without registration or notice, Harvard may have a protectible copyright in the record for up to five years, if it registers the record within that time and makes other good faith efforts to provide notice that it is in fact copyrighted. The MARC records provided by the Library of Congress are not copyrighted because they are works produced by the United States government.

When the Harvard Law Library modifies a record originally input by Yale and attaches its holding symbol to it, to make a record for its own catalog, Harvard has created a derivative work, one "based upon one or more preexisting works," if its modifications are more than trivial variations. If Yale had copyrighted the work and given the requisite notice, then Harvard's use without permission infringes Yale's right in the work.

Unless Yale has given RLIN exclusive rights in its work, RLIN's rights are in no way infringed by this use of one record. However, if Harvard is using a number of records from several different libraries, then it would seem that Harvard may be infringing not only the individual records, but also the collection of records as such. In essence, Harvard would be using the data base to create its own "data base" on its holding tape.

What rights does RLIN have in its data base collection of copyrighted records? Assuming that it is using them "lawfully," RLIN can copyright the data base as either a collective work or a compilation. The data base is not "published" within the meaning of section 101 since it has not been distributed or offered as an entity to other parties. Thus, RLIN need not attach copyright notice for the copyright to be valid, though it may do so if it wishes. Nor does it need

139. Arguably, Harvard need not attach notice because it has not "published" the work by distributing copies to the public. It does seem however that by entering the work into a data base intended to be used by other libraries Harvard is at least "offering to distribute copies . . . for purposes of further distribution or public display," and thus is publishing within the definition of § 101 and must attach notice to be protected.

140. However, if Harvard takes no steps to copyright its catalog record, persons who use it in reliance upon its lack of copyright notice will incur no liability for actual or statutory damages from their infringing acts committed without notice of the copyright. 17 U.S.C. § 405(b) (1982).

141. See Dewey, supra note 66.

to register the data base with the Register and deposit the necessary identifying information, though again it may do so if it wishes.

If, therefore, Harvard searched for and copied all the records in the data base bearing its own document symbol, it might be infringing RLIN's rights. Under one view, there would be no infringement, even if Harvard used records not originating with Harvard, since it would merely be copying the unprotected underlying components of the compilation, but this view is not easily applied to data bases. CONTU espoused the general mediating principle that the use of one item retrieved from such a work . . . would not under reasonable circumstances merit the attention of the copyright proprietor. Nor would it conceivably constitute infringement of copyright. The retrieval and reduplication of any substantial portion of a data base . . . would likely constitute a duplication of the copyrighted element of the data base and would be an infringement.

The tapes, which may be infringing both other libraries' and RLIN's rights, may be obtained by someone who compiles them and offers the new compilation for sale. If the tapes themselves are infringing copies, then use of them would not be "lawful" and the compiler could not obtain copyright protection for his new compilation. Since the compiler obtained the tapes themselves with the libraries' consent, he could argue that he did obtain them lawfully

143. Unprotected as far as RLIN is concerned. Yale, as the author of the record, could still sue Harvard for infringement.

144. M. Nimmer, supra note 26, at § 3.04.

145. Squires notes that the economic rationale for protecting data base compilations should protect against unauthorized searches since:

Economic damage to the proprietor of a data base is precisely what would result from an unauthorized search . . . .

Nor should the fact that the information retrieved by searching a data base is itself in the public domain mitigate the infringing nature of an unauthorized search. It is the organization of information and the capacity to provide that information in response to a request that constitutes the copyrightable element of a directory in hard copy, and it should not affect the scope of copyright protection that the information contained in a directory is embodied in machine-readable media and retrieved by what is, in effect, an automated indexing system.

Squires, supra note 55, at 232.

146. Final Report, supra note 95, at 42.

147. 17 U.S.C. § 103(a) (1982). As Nimmer summarizes it: "One who receives and exploits infringing material from the original infringer, with notice of the infringing nature of the material, may be held to be acting in concert with the original infringer." M. Nimmer, supra note 26, at § 12.04[A] n.7. Thus, in a suit by West Publishing Co. against LRS for infringement of its key number digest, it was held proper to join CSSC, to whom LRS had transferred its infringing computer tapes. Computer Searching Service Corp. v. Ryan, 439 F.2d 6 (2d Cir. 1971).
and might be able to copyright his data base just as RLIN can—no notice being necessary.

In the second phase of the hypothetical, the compiler expands his compilation by adding tapes obtained from the Government Printing Office (GPO). The GPO data that he adds to this new compilation is not itself copyrightable, because it is a government work. If taken from OCLC, however, it infringes OCLC's compilation rights in the data base just as Harvard's tape probably infringes RLIN's. Copying a substantial portion of IAC's data base without permission would infringe IAC's rights in it. Any systematic use by the compiler of the search results obtained from Lockheed's services would probably infringe Lockheed's rights as exclusive licensee of the data base searched, as well as those of the licensor.148 The library that searches the compiler's data base, as in the third part of the hypothetical, will likewise infringe by regular copying and certainly by selling the results of its search.

This discussion resembles a law school examination response to a ridiculous hypothetical question, as infringement piles on infringement. Harvard is infringing the copyrights of RLIN members and probably RLIN itself, the compiler is infringing OCLC's, IAC's and Lockheed's copyrights as well as being liable for the infringements Harvard passed on with its tapes, and Harvard in turn is infringing the compiler's and Lockheed's copyright by copying search results from the data base.

The idea of the Harvard Law Library as a kingpin of copyright crime may have a certain appeal. In the real world, however, the many unavoidable "shoulds" and "mights" in the analysis above dismay rather than delight institutions that desire to avoid the costs of complexity and uncertainty, as well as the possibility of litigation. For this reason, contracts expressly defining rights to the records being used are made in almost all situations similar to those above.

B. RESULTS UNDER PRIVATE CONTRACTUAL ARRANGEMENTS

Most issues of copyright ownership are worked out between the parties involved as a matter of private contract. This provides a new perspective to the situation above. Harvard, like other libraries of the Research Libraries Group, does not in fact copyright its catalog records. The standard RLIN contract provides that "all data added to the RLIN data base by participants may be used by RLG/RLIN for any purpose without restriction."149 Likewise, the OCLC con-

148. See CRC Systems, Inc., supra note 120, at 185; Final Report, supra note 95, at 41-42; Squires, supra note 55, at 232.
149. A sample agreement is reproduced in Matthews, supra note 75, at 815.
tract provides that "OCLC has full right to the OCLC On-Line Cata-
log," and further requires that a library "can not obtain any parts of
the OCLC On-Line Catalog not theretofore contributed or used by
the Library," and that the Library cannot "sell, transfer or grant use
of copies of the records provided to it by OCLC to any other entity
whatssoever." The RLIN contract expressly gives the library the
right "to obtain a copy of the data it enters or of all records contain-
ing data entered by participant" but prohibits the library from copy-
ing any part of the data base in machine-readable form without
RLIN's written consent. Thus, when Harvard inputs an original
record or modifies another library's record for its own use, it in-
fringes no one's rights, and when it receives tapes of its holding
records from RLIN it is merely realizing an express contractual
right.

The RLIN contract, unlike that of OCLC, contains no restriction
on a library's use of its holdings tape. As a result, the library can
feel fairly confident in providing the tape to a compiler, reasoning
that if RLIN wished it could easily restrict third-party use as does
the OCLC contract, which seems to make clear that GPO could not
transfer its tape to the compiler. Of course, GPO might be able to
persuade OCLC to waive the provision in this case. Aside from that,
the compiler's use of the RLIN libraries' tapes is lawful and cer-
tainly his compilation is copyrightable so far as they are concerned.

The remainder of the problem involves commercial vendors of
information rather than non-profit libraries, parties even more likely

150. Id. at 802-03. The contract provision reproduced was current in late 1979. In
spring 1980, after hearings in which libraries argued against restrictions on their ability
to share their tapes with nonmember institutions, OCLC decided against placing
restrictions or limitations on third-party use of records from the data base "on the
condition that such use is for the benefit of participating libraries, networks and
OCLC." Conaway, OCLC to Allow Third-Party Use of Records, 41 College & Re-

However, in December 1982 OCLC copyrighted its data base to protect it from "a
number of practices by some members and third parties that are detrimental to the
continuing value of the data base." (Advances in technology allowed libraries to ma-
nipulate data bases and create products and services from which OCLC could derive
no income.) Copyright notices were posted with the log-on message which greets
each terminal user as well as on libraries' holding tapes produced by OCLC. Present
contracts between OCLC and members "will be deemed to include licenses under the
copyright to permit the exercise of all rights granted under such contracts." OCLC
will negotiate other uses as desired. See OCLC Seeks Copyright Protection for Its

151. Matthews, supra note 75, at 815-16.

152. Harvard and several other libraries do in fact currently supply their RLIN
tapes to the compiler. In connection with OCLC and third party use, see supra note
150.
to make explicit contractual arrangements to prevent uncertainty about ownership rights and possible infringement. Commercial data base vendors do copyright their data bases and cause a notice of copyright to appear on the terminal when each user signs on.\textsuperscript{153} IAC could sell a copy of its data base to the compiler. More likely, it would grant him the license to use it in exchange for royalties based on his use. He would make a similar licensing arrangement with Lockheed for use of search results from their data base. Such arrangements are standard in the industry.\textsuperscript{154}

The next question concerns the non-profit searcher of the compiler's data base who wants to distribute, or possibly sell, the results of that search. Data base vendors have a standard restriction on use of hard copy search print-outs, allowing the user to make one copy for archival purposes, and no more than six for distribution inside its own institution. If the library distributed the copies more widely or sold them, the data base vendor would view its actions as infringement. Copying of search results in machine-readable form, called "down-loading," is prohibited under the standard search access contract. If libraries are willing to pay more than the standard data base fee, they can individually negotiate contracts allowing broader uses of the search results in exchange for special payments.\textsuperscript{155}

Private contractual arrangements can eliminate the danger of unexpected liability for infringement, but only at a price. In this third situation especially, libraries may not be able to afford to pay for the kind of use they would like to make of the search results. CONTU's analysis of this situation should be recalled: "The issue of how much is enough to constitute a copyright violation would likely entail analysis on a case-by-case basis with considerations of fair use bearing on whether the unauthorized copying of a limited portion of a data base would be held noninfringing."\textsuperscript{156}

Thus far there has been no discussion of the concept of "fair use." But it is, in fact, an essential part of any exploration of copyright law and property rights in information. Described as the necessary escape valve or the second side of copyright, "fair use" refers to a body of judicially developed limits on the exclusive rights of

\textsuperscript{153} Personal communication, Marjorie Hlavah, President, Access Information Corporation, New Mexico (Apr. 26, 1983).

\textsuperscript{154} See CRC Systems, Inc., \textit{supra} note 120, for a brief discussion of data base licensing arrangements.

\textsuperscript{155} Some data base vendors are currently negotiating terms with users covering down-loading (e.g., BIOSIS and the Art Alliance.) This paragraph is based on information from Marjorie Hlavah, \textit{supra} note 153.

\textsuperscript{156} \textit{FINAL REPORT}, \textit{supra} note 95, at 42.
copyright owners, given express statutory recognition for the first time in the Copyright Act of 1976.

The concept of fair use reflects its judicial origins as it inevitably requires balancing the interests of the copyright owner against those of the would-be user of the copyrighted work. An exploration of the nature and application of the doctrine provides the best way in which to examine those competing interests, and the underlying policies which should inform and shape the limits of copyright protection as it applies to library bibliographic data.

C. RESULTS UNDER THE FAIR USE DOCTRINE AS A LIMIT ON COPYRIGHT OWNERSHIP RIGHTS

1. Traditional Rationale for Copyright

The Constitution gives Congress the power to grant copyright to authors “to promote the Progress of Science and useful Arts.” The Supreme Court has said that the primary objective of the United States in conferring copyright lies “in the general benefit by the public from the labor of authors,” for though “the immediate effect of our copyright law is to secure a fair return for an author’s creative labor . . . the ultimate aim is, by this incentive, to stimulate artistic creativity for the public good.” Copyright as a means of promoting creation of works of art or science valuable to society has traditionally been viewed as achieving that effect in two important ways.

First, copyright is said to encourage individual creative effort by protecting the creator’s right in the fruits of his labor. Under this view, the constitutional provision for copyright reflects the framers’ grounding in John Locke’s conception of property: “every man has a property in his own person. . . . The labor of his body and the work of his hands we may say are properly his. . . . It being by him removed from the common state nature placed it in, it hath by his labor something annexed to it that excludes the common right of other men . . .” Protected by copyright, the author can exclude the common right of other men and prevent appropriation of his work by others. That this is a significant incentive to creation is shown by a recent survey of scientific and technical authors, who were asked to rank a list of possible motives for authorship. Their primary objective was found to be to “establish intellectual ownership of the work.”

---

158. Twentieth Century Music Corp. v. Aiken, 422 U.S. 151, 156 (1975).
159. J. LOCKE, SECOND TREATISE ON CIVIL GOVERNMENT ch. 5, ¶ 27 (1789).
160. Boucher, The Place of the Author in the Coming Information Society, in
less claim to new or deep intellectual content presumably feel the desire to establish rights in their work as a product of their intellectual efforts.

Authors seek more than psychic rewards for their labors. Indeed, the rationale most often advanced for copyright protection is that it provides an economic incentive for authors to create works and for publishers to disseminate them. The need for this special incentive is attributed to the unique nature of creative works. The initial cost of their production is very high, but the cost of replication is generally very small. Unless the replicator can be made to bear part of the cost of the production, no private person will be willing to make the large investment that it requires. Copyright is a way of preventing replication against the creator's will, giving him the ability to create knowing that his efforts will be rewarded.

As the Supreme Court expressed it, "the encouragement of individual effort by private gain is the best way to advance public welfare through the talents of authors." Courts and commentators stress that the primary objective of copyright is not the welfare of the individual author or publisher but the advancement of public good which flows from creation and dissemination of new works. The importance of a free flow of information to societal welfare is unquestioned. "[A]ll human organizations, no matter how simple, depend for their functioning on an intangible resource called 'information' . . . the more complex the society, the more central information is to its economic activities." As well as serving the economic activities of society, information can "stir intelligence and sensibilities" and is vital to the democratic process.

---

STRATEGIES FOR MEETING THE INFORMATION NEEDS OF SOCIETY IN THE YEAR 2000 at 117, 120 (M. Boaz ed. 1980) [hereinafter cited as Boaz].

161. Even scientific publishers listed revenue as the most important reason for their activity. Boucher, supra note 160, at 125. See also R. Saltman, Copyright in Computer-Readable Works: Policy Impacts of Technological Change 12 (1977).

162. For one explanation of the economic rationale for copyright, see Braunstein, Economics of Property Rights as Applied to Computer Software and Data Bases, in TECH. AND COPYRIGHT, supra note 55, at 233, 237-38.


164. See Twentieth Century Music Corp. v. Aiken, 422 U.S. 151, 156; B. Kaplan, An Unhurried View of Copyright '74 (1966); Note, Toward a Unified Theory of Copyright Infringement for an Advanced Technological Era, 96 Harv. L. Rev. 450, 462 (1982).


166. B. Kaplan, supra note 164, at 74.

What has been questioned, however, is what role copyright really plays in fostering art and scientific knowledge. Stephen Breyer (now Judge Breyer of the First Circuit) challenged the traditional premises of copyright law by asserting that it neither served the "moral" interests of the author nor gave an indispensable economic incentive to authors and publishers. He first argued that the author's personal interest in his creation was not in itself a sufficient reason for making it his property; property rights have no necessary relation to labor expended. Further, the author's interest in preventing misuse (by deforming alteration, unauthorized publication, or misattribution of his work) could be better protected by tort remedies.\(^{168}\)

Turning to the economic rationale, he noted that there was no clear relation between copyright and the amount of compensation authors received:

Simply to speak of "fruits of one's labor" does not show that the author should be paid more than his persuasion cost or how much more he should be paid. In particular, it does not demonstrate that the amount he receives under existing copyright law is any more "just" than what he would receive without copyright protection or under a different copyright system . . . . Nor has copyright proved particularly effective in providing compensation for the "great work" of unusual social value.\(^{169}\)

The central portion of the article is devoted to an analysis of the probable effect elimination of copyright would have on publishers of books. Breyer concludes that "to abolish protection would not produce a very large or a very harmful decline in most kinds of book production. Abolition should benefit some readers by producing lower prices, eliminating the cost of securing permission to copy, and increasing the circulation of the vast majority of books that would continue to be produced."\(^{170}\) Though uncertainty as to the effect of abolition of copyright and its admitted negative effects on some classes of books led Breyer to eschew advocating abolition, he did conclude that the then-pending Copyright Act should not contain any stronger provisions for protection.

Breyer's economic analysis has been challenged.\(^{171}\) However,
his general characterization of the "case for copyright" as "uneasy" has been echoed in many recent discussions of copyright in connection with technological developments. Focusing on the unprecedented central role information has come to play in our society, commentators note the need for adjustment in a legal system which in many respects has not yet been modified to reflect that change.\textsuperscript{172}

2. \textit{Changes That May Undermine the Traditional Rationale for Copyright}

Discussions of the information society reflect Daniel Bell's description of the United States as a "post-industrial" nation: "a post-industrial society is organized around information and utilization of information in complex systems, and the use of that information as a way of guiding the society."\textsuperscript{173} An Office of Technology Assessment study gives some of the indications of this transformed society: the information sector now accounts for over one-half the United States work force, with the service sector of the economy growing at a relatively faster rate than the industrial or agricultural sectors.\textsuperscript{174} "Access to information in the postindustrial society is the key to all social roles—work, citizenship, recreation, and others."\textsuperscript{175}

Not surprisingly, "the value of information increases as society becomes more dependent on it."\textsuperscript{176} As the president of the Information Industry Association described it:

\begin{quote}
The information industry is best understood in terms of information as a commodity, or an economic good, like furniture, shoes, food stuffs, liquid oxygen, etc. The industry structure is like the structure of any other industry. There are producers (of data bases or market research studies, for example), distributors (Lockheed, SDC, BRS), retailers (FIND, Information on Demand), and systems companies capable of assembling the pieces to make up the previous three sectors. An exhaustive survey of the information industry places its total annual U.S. sales at $9.4 billion in 1979. Yes, information is a commodity.\textsuperscript{177}

In some ways, information is unlike other commodities: it is re-
\end{quote}

\textsuperscript{172} See, e.g., B. Kaplan, supra note 164, at 75.
\textsuperscript{173} See, e.g., Weingarten, \textit{The Financing and Governance of Information Networks of the Future: The Public Sector}, in Boaz, supra note 160, at 163, 166.
\textsuperscript{174} Remarks by Daniel Bell before the House Committee on Science and Astronautics, quoted in R. Saltman, supra note 161, at 52.
\textsuperscript{175} O.T.A., supra note 165, at 12, 47.
\textsuperscript{177} O.T.A., supra note 165, at 49.
producible, theft does not take it away from the original owner, and its replication cost is very low in comparison with its original creation cost. Thus, the value of information is not additive. Two copies of the same item are not worth much more to the possessor than one. Its lifetime can be very brief since it can be communicated instantaneously.\textsuperscript{178}

The sheer magnitude of information also sets it apart from other commodities. Information has been exploding exponentially. As one researcher's findings were summarized:

[T]he total number of scientific and technical articles published in scholarly journals in the United States advanced from 105,932 in 1960 to 155,345 in 1975. The number of new books in hardcover form published in the field of science increased more than 40\% in a ten-year period. The number of patents issued in the United States in the period between 1960 and 1974 was 60\% of the total number of U.S. patents issued through 1960. These figures represent only the growth of primary literature. The secondary literature has been growing rapidly, too, with figures from one abstracting and indexing service showing a dramatic increase in the number of abstracts and indexes issued.\textsuperscript{179}

This explosion of information is made possible by changing technology in the production of information, as computerized typesetting, on-line editing, and digital communications increase production of print forms and forms storable in computerized form or video discs. Distribution has been changed by the advent of new data communication systems. Bibliographic data identifying the mass of information is now prepared on computers and is generally offered as a database to searchers.\textsuperscript{180}

This technological change affects traditional concepts of ownership. As the author of the appropriately titled \textit{Goodbye Gutenberg} put it:

Knowledge has been "owned" by the institutions and individuals who have created it, but in the electronic phase of knowledge storage and retrieval, information goes back to society or mankind in general. Power over knowledge remains with the institutions that control the systems by which it is transmitted and obtained, but information can move endlessly from place to place, from storage to storage, without being "copied" or "learned". The growing sovereignty of the audience over the selection of information occurs simultaneously with the rise of new controlling powers over databases. Thus, the triad of publisher, author, public, is trans-

\textsuperscript{178} O.T.A., supra note 165, at 49.
\textsuperscript{179} Boaz, supra note 175, at 12 (summarizing F. Lancaster, Toward Paperless Information Systems 66-103 (1978)).
\textsuperscript{180} O.T.A., supra note 165, at 163-66.
formed into a new triad of database controller, information collector/provider, reader.\textsuperscript{181}

Technology may change the concept of an author from one who autonomously creates a new work to one whose creativity lies in selection and reformulation of information from a vast existing (and available) store.\textsuperscript{182} It has already "undermined the ability of copyright owners to control the distribution of their works; publishing has become easy, inexpensive and anonymous."\textsuperscript{183} Not surprisingly, some commentators have concluded that application of copyright, based on the concept of the autonomous author and the readily identifiable publisher who invests substantial labor in publication, has become inadequate or inappropriate for present-day situations.\textsuperscript{184}

\begin{itemize}
\item \textsuperscript{181} A. Smith, \textit{Goodbye Gutenberg: The Newspaper Revolution of the 1980's} 313 (1980).
\item \textsuperscript{182} Id. at 315.
\item \textsuperscript{183} Note, \textit{supra} note 164, at 451.
\item \textsuperscript{184} Some critics emphasize existing alternate means of protection for legal rights in information products. In § 301 of the 1976 Act Congress provided for federal preemption:

\begin{quote}
[A]ll legal or equitable rights that are equivalent to any of the exclusive rights within the general scope of copyright as specified by section 106 in works of authorship that are fixed in a tangible medium of expression and come within the subject matter of copyright as specified by sections 102 and 103... are governed exclusively by this title... no person is entitled to any such right or equivalent right in any such work under the common law or statutes of any State. 17 U.S.C. § 301(a) (1982).
\end{quote}

The federal statute does not preempt common or state law remedies for subject matter outside the bounds of §§ 102 or 103, or remedies for "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." 17 U.S.C. § 301(b)(3) (1982).

An early draft of the section listed certain rights which would not be equivalent to the preempted rights, including misappropriation and "breaches of contract, breaches of trust, trespass, conversion, invasion of privacy, defamation and deceptive trade practices, such as passing off and false representation," S. 22, 94th Cong., 1st Sess. § 301 (1975).

The Justice Department objected to the use of "misappropriation" and the entire section was deleted. Still, the other causes of action would almost certainly not be preempted. Bigelow, \textit{supra} note 10, at 8. A discussion of these causes of action is unfortunately outside the scope of this Article. For a brief description of them, see \textit{Final Report, supra} note 95, at 16-18.

Besides state law remedies for violation of proprietary rights, a common suggestion is that such rights be limited to the right to receive reimbursement upon use. The idea of compulsory licensing will be explored in the context of bibliographic utilities \textit{infra} subsection 4.
3. Fair Use As Paradigm of the Existing System
   a. Fair Use and the Copyright System

Both critics and defenders of the present copyright system tend to focus on the doctrine of fair use, a judicially developed concept which allows reasonable uses of copyrighted materials under certain circumstances. Sometimes described as a safety valve, it is employed to avoid "unduly restricting the dissemination and use of copyrighted works," which would defeat the copyright statute's purpose.185

The Copyright Act of 1976 recognized this doctrine, restating it in section 107:

Notwithstanding the provisions of section 106, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall 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.186

Congress intended to "restate the present judicial doctrine of fair use, not to change, narrow or enlarge it in any way."187 It further described the nature of the doctrine:

The endless variety of situations and combinations of circumstances that can arise in particular cases precludes the formulation of exact rules in the statute. The bill endorses the purpose and general scope of the judicial doctrine of fair use, but there is no disposition to freeze the doctrine in the statute, especially during a period of rapid technological change. Beyond a very broad statutory explanation of what fair use is and some of the criteria applicable to it, the courts must be free to adapt the doctrine to particular situations on a case-by-case basis.188

Courts both before and after the statutory embodiment of the

188. Id.
doctrine have felt "free to adapt" it, and while the statement of standards themselves has varied little the flexibility of their application has led to characterization of the fair use issue as "the most troublesome in the whole law of copyright."

Fair use is a flexible mediating device between the competing interests of the copyright owner and the would-be user (and society at large), but this is viewed by critics as typical of the inadequacy of the copyright system to deal with new technology. In support, critics point to cases applying the fair use criteria that have reached contradictory results on analogous facts. Fair use can be defended, however, as exactly that aspect of the copyright system which makes possible its just application to new technologically created situations.

In order to examine whether fair use enables the copyright system to work in a context involving a new technological development, it will be necessary to return to the library situation described above. The results reached by the actors through private agreements have been compared to the ownership rights conferred upon them by statutory copyright law. When the fair use doctrine is applied as a limit on those statutory rights, do the results approach those reached by the parties themselves (presumably the desired ones), and if not, are the results nevertheless defensible? This writer believes that fair use, in the library situation at least, does enable existing copyright law to provide a satisfactory framework for resolution of ownership rights in bibliographic data, though not necessarily the best or only one.

b. Results In Library Applications
   i. Use of a Library's Original Record by Other Libraries

The first problem in our hypothetical was presented by Harvard's use of another library's original record. If that original record was copyrighted, could Harvard defend its adoption of the

189. See Note, supra note 164, at 455.
190. Dellar v. Samuel Goldwyn, Inc., 104 F.2d 661, 662 (2d Cir. 1939). The Supreme Court most recently applied the fair use doctrine in Sony Corp. of America v. Universal City Studios, 44 S. Ct. Bull. B641 (Jan. 17, 1984). There it held a manufacturer of video tape recorders was not liable as a contributory infringer of copyrights in programs broadcast on public airwaves. The decision was based in part on the grounds that users of video tape recorders engaged in noncommercial activity had little demonstrable effect on the potential market for the copyrighted works. Id.
OWNERSHIP OF ACCESS INFORMATION

record as fair use? Application of the four criteria for fair use defined in the statute leads to the conclusion that it could.

The first factor is the amount and substantiality of the portion used in relation to the copyrighted work as a whole. Harvard has used substantially all of the copyrighted work, a fact that in most cases would militate against a finding of fair use.\footnote{192} Schroeder v. William Morrow & Co., 566 F.2d 3, 6 (7th Cir. 1977) (court held that where material taken from plaintiff's directory was substantial in quantity, defendant's use would not be fair use). See also Walt Disney Productions v. Air Pirates, 581 F.2d 751 (9th Cir.), cert. denied, 439 U.S. 1132 (1978) (fair use defense could not be applied to almost complete copying of cartoon character); Meredith Corp. v. Harper & Row, Inc., 378 F. Supp. 686 (S.D.N.Y.), aff'd, 500 F.2d 1221 (2d Cir. 1974) (copying of substantial portions of competing textbook held not fair use). But see Sony v. Universal, 44 S. Ct. Bull. (CCH) at B673 (although complete copying ordinarily mitigates against a finding of fair use, the weight of other factors can overcome that general principle); Williams v. Wilkins Co. v. United States, 487 F.2d 1345 (Ct. Cl. 1973), aff'd per curiam by an equally divided court, 420 U.S. 376 (1975) (library photocopying of an entire work can still be fair use).

More difficult to evaluate is the second factor, the effect of the use upon the potential market for or value of the copyrighted work. In effect, Harvard may have simply decreased the potential users/purchasers of Yale's catalog record by one. Alternatively, one could say that Harvard has set itself up as a competing source of the catalog record, which could completely eliminate Yale's market, since Harvard did not have the expense of creating the record and could recover its costs at a lower price than could Yale.\footnote{193} Courts have recognized the difficulty of providing concrete proof of economic harm in infringement cases. Some have allowed a copyright owner to prove economic harm by showing only that the probable economic effect of the use will be harmful, a position partially reiterated by the Supreme Court in Sony v. Universal.\footnote{194} Yale might well be able to make this type of showing.

The third factor, nature of the work copied, has not been of crucial significance in many cases.\footnote{195} In this case, however, the nature

\footnote{192.} Schroeder v. William Morrow & Co., 566 F.2d 3, 6 (7th Cir. 1977) (court held that where material taken from plaintiff's directory was substantial in quantity, defendant's use would not be fair use). See also Walt Disney Productions v. Air Pirates, 581 F.2d 751 (9th Cir.), cert. denied, 439 U.S. 1132 (1978) (fair use defense could not be applied to almost complete copying of cartoon character); Meredith Corp. v. Harper & Row, Inc., 378 F. Supp. 686 (S.D.N.Y.), aff'd, 500 F.2d 1221 (2d Cir. 1974) (copying of substantial portions of competing textbook held not fair use). But see Sony v. Universal, 44 S. Ct. Bull. (CCH) at B673 (although complete copying ordinarily mitigates against a finding of fair use, the weight of other factors can overcome that general principle); Williams v. Wilkins Co. v. United States, 487 F.2d 1345 (Ct. Cl. 1973), aff'd per curiam by an equally divided court, 420 U.S. 376 (1975) (library photocopying of an entire work can still be fair use).

\footnote{193.} See Freid, supra note 185, at 485. The price that could be charged is limited by the amount other libraries are willing to pay, which will be determined by the cost of original cataloging for each library.

\footnote{194.} 44 S. Ct. Bull. (CCH) at B674. In Sony the Supreme Court stated that "[w]hat is necessary is a showing by a preponderance of the evidence that some meaningful likelihood of future harm exists. If the intended use is for commercial gain, that likelihood may be presumed. But if it is for a noncommercial purpose, the likelihood must be demonstrated." Id. See Freid, supra note 185, at 473 for a discussion and collection of other cases.

\footnote{195.} Id. at 467. Some have suggested that the scope of the fair use protection will depend on the nature of the work, with greater license to use a "work more of diligence than of originality or invention." New York Times Co. v. Roxbury Data Interface, Inc., 434 F. Supp. 217 (D.N.J. 1977). Cf. MCA, Inc. v. Wilson, 677 F.2d 180, 182 (2d
of the work is significant, as it is intimately tied up with the purpose of the use.

The fourth factor has been viewed as perhaps the most important determinant of fair use. Courts have tended to allow fair use broadly where the use is connected with "the field of learning" but narrowly "where taking is solely for commercial gain." Harvard could make the catalog record available to other libraries for a price, but its primary purpose in copying the record is to use it in its catalog, indicating to users that the work is part of its collection.

The library catalog is certainly connected with the process of research and learning. However, it does not fit comfortably within the terms courts generally use to describe permissible purposes of fair use, such as "advancement of science." In one view, free dissemination of bibliographic data does less to advance knowledge than does, for example, data embodying results of medical research, for it is "intended to be used as a guide to something further and different." Yet in another sense bibliographic data should be the most freely disseminated information of all, since it is used as part of the process of research to create new knowledge. Given the overall purpose of the library catalog in development of information, courts are very likely to find that one library's use of another's catalog record is indeed a fair use. The beneficial purpose of the use outweighs the complete copying (for the purpose of substituting rather than supplementing the original) and possible economic effect which ordina-

---

196. Freid, supra note 185, at 468.
198. "Advancement of science" is the rubric often used to describe the paradigmatic use protected by the fair use doctrine. See, e.g., Loew's, 131 F. Supp. at 193; Thompson v. Gernsback, 94 F. Supp. 453, 454 (S.D.N.Y. 1950); Towle v. Ross, 32 F. Supp. 125, 127 n.4 (D. Or. 1940). One court described the works that might advance science: "The classic instances in which the courts, under the aegis of 'fair use,' have permitted authors to use excerpts from a copyrighted work . . . are those of literary criticism and parody of the copyrighted work, history and biography," H.C. Wainwright & Co. v. Wall Street Transcript Corp., 418 F. Supp. 620, 625 (S.D.N.Y. 1976).
rily would allow the plaintiff to prevail.\textsuperscript{200}

This result, that libraries should be able to copy cataloging data freely, is one that libraries have reached on their own. Consistent with the library's mission of making information available to all, co-operative exchange of bibliographic data is a long-standing library tradition.\textsuperscript{201} Bibliographic data is indeed a valuable commodity whose production is costly and should be encouraged by giving an economic incentive to its creator. However, knowledge of a library's resources is also a public good that should be encouraged, and fair use in this instance promotes the copyright statute's aim of encouraging free dissemination of information for the public good.

It should be acknowledged, however, that copying of a catalog record does not fit well into the traditional area of fair use. Fair use has traditionally encompassed the use of the original for a different purpose, such as a critical commentary on the original, or what one commentator has called the "derivative" function.\textsuperscript{202} Rather, here the use is iterative—an exact duplication for the same purpose.\textsuperscript{203} Still, in this rather unusual situation of copying a catalog record, the fair use doctrine is able to reach a functional and defensible result. Though the result is proper—upholding the right to disseminate bibliographic data in non-profit situations—it comes at the expense of removing Yale's incentive to produce the original data.\textsuperscript{204}

\textsuperscript{200} There appear to be no cases on the question of fair use in connection with a bibliographic entry as such, and none in connection with library catalog records. The closest factual analogy would be the case holding individual trade catalog descriptions protectable. \textit{See supra} text accompanying note 38. However, there the defendant was a for-profit operation and the question of fair use was not litigated.

\textsuperscript{201} \textit{See supra} § IV(A)(1).

\textsuperscript{202} \textit{See Note, supra} note 164, at 462.

\textsuperscript{203} One can argue that if Harvard does not in fact use the record in any profit-making activity, there is little real detriment to Yale. \textit{Note, supra} note 164, at 465-66. It does seem that Harvard has taken an unfair advantage of Yale's labor. Here of course the de minimis factor supports the fair use conclusion. Yale is very unlikely to complain of copying of one of its records. Were Harvard to copy the whole catalog the result might be different, as the “character of the use” would be more unfair. Were a commercial vendor to do so, the "use" factor would almost certainly lead to a finding of infringement because of the commercial nature of the use. \textit{See} H.C. Wainwright & Co. v. Wall Street Transcript Corp., 418 F. Supp. 620 (S.D.N.Y. 1976).

\textsuperscript{204} It can be argued that in the Harvard-Yale situation Yale's incentive is not reduced, because it knows that it will be able to obtain Harvard's original records in exchange for its own in approximately equal numbers. However, many members of bibliographic utilities do little original cataloging, sometimes waiting for several months until they find cataloging in the database for their new acquisitions. The problem of these "net takers" or "freeloaders" has aroused concern by other members. In May, 1982, an OCLC Users Council Special Task Force on Responsible Use of the OCLC System recommended that a code of ethics for OCLC users be developed, to try to stem the "freeloading" problem. \textit{Bibliographic Utilities, supra} note 71, at
More satisfactory is the solution the parties themselves have reached in connection with the bibliographic utilities. In RLIN, libraries who use already-existing records from the data base pay a fee for each record. Libraries who input original records are charged nothing for the cataloging support service provided by the utility for that record. Thus, libraries essentially do pay each other for use of original data. Fair use, in contrast, is all or nothing. If the use is fair, there is no infringement and therefore no liability for payment.

As between libraries, application of the fair use doctrine produces a result that is acceptable, if not optimal. The next case, that of the compiler who uses the RLIN tapes of several libraries, presents more difficulty.

ii. Compiler’s Use of Library Bibliographic Utility Holding Tapes

Even if a compiler obtains tapes from most of RLIN’s members, for the purpose of current information dissemination the tapes used will cover only holdings added during a relatively short period of time—thus the portion of RLIN’s data base used is likely to be insubstantial. The effect of the compiler’s use on the potential market for or value of RLIN’s data base is debatable. On the one hand, use of the tapes as part of an SDI service does not compete with RLIN’s economic rights in its data base as a cataloging utility; the compiler is not directly competing with RLIN by setting himself up for the same purpose. Courts may therefore be more likely to find this non-competing use a fair one.

On the other hand, the compiler has gravely decreased, if not...
eliminated, the potential market RLIN might have if it decided to use or license use of its data base for SDI services. Courts have split on the implications such an effect has for fair use.208

At least one case has held that this "opportunity loss" does not prevent a finding of fair use. New York Times v. Roxbury Data Interface209 involved a suit by the New York Times for infringement of its annual index to the newspaper. The defendants had compiled a personal name index to the annual index, so that a user could look in one place to find references to all the volumes of the New York Times Index in which a particular name appeared. The defendants did not key their index to the pages of the newspaper itself—thus it was still necessary to use the New York Times Index to find the location of articles about the named person. The court held the defendants' use of the New York Times Index to be fair use.

One of the plaintiffs' contentions was that the defendants' work had a detrimental effect upon plaintiffs' potential market, since it would adversely affect any personal name index they might publish in the future.210 The court remarked that "nothing stands in the way of plaintiffs' publishing a work to rival defendant's index."211 As the court viewed the argument, plaintiffs were attempting to portray themselves as authors of a book who should be able to prevent the unauthorized making of a derivative work (such as the personal name index) derived from their work. The court rejected that analogy, finding that the defendants had not copied the substance of the plaintiffs' work (that is, the page cites to the newspaper), that defendants' work had a different function, and that no creative idea of plaintiffs' was appropriated by defendants.212

Of course, in Roxbury Data, the use of the defendants' index required use of plaintiffs' index to reach the newspaper articles sought, thus making the plaintiffs' market effect argument very weak. The case can therefore be squared with the seemingly contrary case law holding that "if defendant's work adversely affects the value of any of the rights in the copyrighted work . . . the use is not fair even if the rights thus affected have not as yet been exercised,"213 for it can be said that here the holding simply rested on the fact that there was no adverse effect on plaintiffs' rights to pro-

208. See M. Nimmer, supra note 26, at § 13.05(B), for a full discussion and collection of the cases.
210. Id. at 224.
211. Id.
212. Id. at 224-25.
213. M. Nimmer, supra note 26, at § 13.05(B), citing Iowa State University Research Foundation, Inc. v. American BroadcastingCos., 621 F.2d 57 (2d Cir. 1980).
duce a personal name index. The *Roxbury Data* result therefore may not be a bar to RLIN in a suit against the compiler if it could show that its rights in future SDI use of the data base had really been affected.

Even lacking such a showing, RLIN could find support in *National Business Lists v. Dun & Bradstreet, Inc.* Plaintiff in the case took names of and data about business executives from the defendant's directory, entered the material in a data base, and used it to generate mailing lists, which it sold. Apparently Dun & Bradstreet also sold, or planned to sell, mailing lists based on its directory. Although the court ultimately held that the defendant was largely estopped from invoking copyright protection because it had acquiesced for many years in the plaintiff's use, it examined plaintiff's claim of a fair use privilege at some length, and affirmed the jury's finding of infringement beyond fair use. In so doing, it distinguished *Roxbury Data*, where the defendants' index was likely to have little effect on the plaintiff's market and where the defendant had expended substantial effort in producing the derivative work. Although it could therefore have disregarded *Roxbury Data*, the court nevertheless discussed the *Roxbury Data* reasoning, and criticized that court's disregard of *Leon v. Pacific Telephone & Telegraph Co.*

As was the case in *Leon*, the defendant's directory in *Roxbury Data* "was different in arrangement, served a different purpose, and was only potentially competing." These factors did not preclude a finding of infringement in *National Business Lists* though in *Roxbury Data* they did. The *National Business Lists* court felt that *Leon*'s finding of infringement based on the economic factor was correct. It noted that since the information was stored in machine-readable form that could be produced as output in varying formats "an emphasis upon arrangement and form in compilation protection becomes even more meaningless than in the past." The economic factor of labor expended by plaintiff and avoided by defendant thus becomes the crucial determinant of protection. As the court described it, there are "notions of unfair competition implicit in copyright protection of diligent application." The court suggested that:

---

215. *National Business Lists* sued Dun & Bradstreet for antitrust violations; Dun & Bradstreet countered with the copyright infringement claim.
216. 91 F.2d 484 (9th Cir. 1937).
218. *Id.* at 97.
219. *Id.* at 95.
OWNERSHIP OF ACCESS INFORMATION

The directory cases . . . are the most striking illustration in copyright law that the misappropriation doctrine . . . has there long found a house, if not a home. Compilations, being more the product of diligent application and less the result of intellectual creativity than possibly any other form of protectable work, are at one end of a spectrum extending to art, poetry and music. Diligent application has, through copyright, been accorded a measure of protection because that is the only protection which is meaningful.220

The court accordingly described the appropriate test of economic harm in a compilation case: "Concepts of unfair competition presuppose an assessment of the 'copyist's' behavior in light of the competitive relationship he has to the copyright owner or at least in light of the impact upon the owner's legitimate expectations of business advantage."221 Under the National Business Lists approach, although the compiler is not in a directly "competitive relationship" to RLIN because his compilation is for a different purpose, it may have an impact upon RLIN's "legitimate expectations of business advantage" and thus can be found infringing.

In contrast, the general language of Roxbury Data suggests that the compiler would prevail against RLIN on the issue of "market effect." This is especially true given the line of cases Roxbury Data cites which recognize that noncompetitive uses may be more entitled to a finding of fair use than those which directly compete.222

It is not at all clear what result a court might reach as far as the criterion of market effect is concerned. Even the National Business Lists court did not view the economic effect as determinative of fair use by itself, though it might establish a prima facie case of infringement.223 The two remaining factors must therefore be examined: nature of the work and character of the use.

The nature of the copyrighted work can be described in two ways—either as knowledge that should be freely available, or as a valuable arrangement that should not be infringed. As noted above, a number of courts have held that fair use is more broadly construed when the copyrighted work is a compilation than when it is a work of more originality.224 Although there are no cases concerned with fair use of bibliographies, a number of cases involving law di-

220. Id.
221. Id.
222. Roxbury Data, 434 F. Supp. at 223.
223. "[T]he concept [of fair use] extends well beyond the unfair competition aspects of diligent application cases . . . ." National Business Lists, 552 F. Supp. at 96. The court rested its finding of infringement partly on the substantial nature of the copying as well as the commercial nature of the use. Id. at 96-97.
224. See supra note 189.
gests and treatises draw a distinction between the permissible use of citations and the impermissible use of text.

Where defendant compilers of a digest used a list of cases from plaintiff's legal encyclopedia, the court found the use permissible. "If it be held that an author cannot consult the authorities collected by his predecessors, the law of copyright, enacted to promote the progress of science and the useful arts, will retard that progress."[225] The court stressed that the defendant had copied none of the plaintiff's text, and had independently examined the cited cases before inclusion. Another law book case held similarly that "a subsequent author may use and copy the citations of a prior author, provided he examines and verifies such cases before using them."[226]

In contrast is the slightly later case of *W.H. Anderson v. Baldwin Law Pub. Co.*[227] A one-volume treatise on the Ohio Code was held to infringe the plaintiff's multi-volume work on the same subject. The court rejected the argument that independent checking of cases copied from the plaintiff's list was enough to protect the defendant: "[E]ven if those cases were subsequently read and independently abstracted, the labor of finding all Ohio cases construing statutes and deciding under what section of the Code to deal with them was thus saved to defendant."[228]

The different results illustrate the conflicting conceptions of bibliographic data. On the one hand, bibliographic data is viewed as information in the public domain that can be freely used. On the other hand, it is viewed as a valuable compilation, which should be protectible as the fruits of labor. Thus, even in dealing with identical data courts can reach different conclusions based on the "nature of the work." One commentator who tried to summarize the "confused" law book cases concluded that: "it may be that the basic issue in each case is whether an earlier work has been collaterally used or substantially copied as well."[229] Although this is not entirely accurate (the holding in question rested more on an "unfair competition" notion than on one of "substantial copying"), it does point up the importance of the fourth factor, the character of the use.

The purpose and character of the use in the example may be

---

226. White v. Bender, 185 F. 921, 925 (S.D.N.Y. 1911).
227. 27 F.2d 82 (6th Cir. 1928).
228. 27 F.2d at 88.
the factor that will outweigh the others and lead to the conclusion that the copyright has been infringed by the compiler. RLIN is a non-profit organization supported by the fees of the members who provide its data. Suppose the compiler is going to use portions of that data base for profit. It is true that “the fact that defendants seek to profit financially will not preclude that use from being a fair use.” Nevertheless, courts have generally found commercial use a factor that weighs heavily against fair use. One reason may be the notion that if the defendant will make a profit from his use he can afford either to pay the copyright owner for that use, or to pay for the original labor necessary to avoid infringement of the copyright. More generally, courts reason that publications for profit do not fall within the category of scholarly research. Use of material in them therefore cannot be justified as fair use for the purpose of the doctrine, which is to further the progress of art and science by preventing undue restrictions on the use of copyrighted material in new works of learning.

The problem with the distinction between commercial and educational uses is that it can be a difficult one to make. As the Supreme Court said in another context, “the line between the transmission of ideas and mere entertainment is too elusive for this court to draw, if indeed such a line can be drawn at all.” One commentator characterizes the purpose test as “bankrupt” and notes that the commercial use factor in any event no longer precludes a finding of fair use.

Timberg would redefine the fairness test to make the nature of the user rather than the use determine whether a privilege was available. He advocates a broad fair use privilege for those in professions advancing the arts and sciences who “in good faith carry out dissemination of information, ideas and criticism to the public.” His examples include “teachers, researchers and scholars.” While his broad language might cover the compiler's use as “dissemination of information” he would not include the compiler in the group of

---

230. *Roxbury Data,* 434 F. Supp. at 221. The *National Business Lists* court agreed with the view that commercial character of use did not per se preclude a finding of fair use. 552 F. Supp. at 96.

231. See cases cited *supra* note 197.

232. See cases cited *supra* note 185.


234. See *Roxbury Data,* 434 F. Supp. at 221; *National Business Lists,* 552 F. Supp. at 96. Timberg quotes the Second Circuit to that effect: “Whether an author or publisher has a commercial motive or writes in a popular style is irrelevant to a determination of whether a particular use of copyrighted material in a work which offers some benefit to the public constitutes a fair use.” *Rosemont Enter. v. Random House, Inc.,* 566 F.2d 303, 307 (2d Cir. 1977), quoted in Timberg, *supra* note 191, at 220.
those to whom the privilege should automatically be granted.\textsuperscript{235} Although the nature of the user may in many cases be a helpful test of whether the use is "fair," for a commercial user one would still have to examine the other factors, as in some cases such a user might be entitled to a fair use privilege. Further, the test fails to address the situation in which use by a non-profit organization is heavily commercial.

A better approach would be to consider the nature of the user together with that of the use. In the hypothetical situation this Article presents the user is a for-profit entity. However, it is using tapes made available by individual libraries. Presumably at least some of the members of the non-profit RLIN organization wish to have the compilation, viewing it as a public service. Still, the user is charging the libraries for its manipulation of the data they provide. By either a "use" or "user" test, the compilation appears to fall outside the accepted bounds of fair use.

It has been shown that opposing views of the nature and character of the use and its economic effect on RLIN are possible. A court weighing the four factors would have a difficult case and could easily and justifiably reach either a finding of fair use or one of infringement.

Interestingly, the utilities themselves have reached different conclusions. RLIN imposes no restrictions on the use a library may make of its tapes. The Harvard Law Library has even given its tapes to a compiler. In contrast, OCLC in 1982 copyrighted its data base, expressing continuing concern about unauthorized third party use.\textsuperscript{236} The tensions between the library's rights in its cataloging records and the data base proprietor's rights in those records are manifest in this situation. It is not surprising that the law provides no clear resolution of the issues when the parties themselves disagree as to acceptable solutions.

\textsuperscript{235} As noted above, Timberg would separate the question of whether a fair use privilege exists from whether the user had an obligation to pay. A user whose operation was for-profit would have the obligation to pay for its use. Under Timberg's approach, the compiler might have a privilege to use RLIN's data base, but would have to pay RLIN for it. Timberg, supra note 191, at 221.

\textsuperscript{236} It appears that under the standard OCLC contract a library could not take a tape of its holdings from the OCLC data base over OCLC's objections. As a practical matter, many libraries lack the technical capacity to obtain such information without OCLC's assistance, but technological developments make such uses more possible. OCLC's copyrighting of its data base was partly a reaction to such potential uses. See supra note 150.
iii. Reproduction of Search Print-outs from Commercial Data Bases

Before discussing the question of whether such ambiguity is a sign of intolerable confusion or healthy flexibility, the final hypothetical should be examined. In this third example, the compiler has performed a search for the library and the library distributes printed copies of the search results to the faculty each month, or sells them for a small fee to students, or makes them available to the general public at a profitable price.

Library use, reproduction, and distribution of part of a copyrighted work is the situation to which most commentary about the fair use doctrine is directed. For many, the question of "fair use" is synonymous with the question of the right of libraries to photocopy parts of publishers' printed works. Congressional hearings and much of CONTU's work focused on that issue. The result was a special statutory provision allowing certain types of photocopying by libraries as a limit on exclusive ownership rights.

In literal terms, section 108 refers to a library's right to "reproduce" a "copy" of a "work" and could arguably be applied to the copying of portions of a data base. In fact, Congress intended the section to address little more than the issue of photocopying. Making copies of a search print-out does not fit within any of the definitions of non-infringing uses. In fact, the situation described above would seem to fall within "concerted reproduction or distribution of multiple copies or phonorecords of the same material, whether made on one occasion or over a period of time, and whether intended for aggregate use by one or more individuals or for separate use by the individual members of the group," which is explicitly not allowed.

The fact that Congress did make special provision for library use (which would not necessarily be within the general definition of fair use) can be interpreted as a recognition that libraries serve a special function in the dissemination of information, and that their public utility outweighs considerations of copyright. However, the carefully drawn limits on library rights suggest that Congress was unwilling to carry that interest in public dissemination too far. Section 108, then, is of little real assistance in approaching the issue of copying of data base search results by libraries. Therefore, it is necessary to turn to the factors defining fair use under the statute.

The first, the nature of the copyrighted work in the hypothetical situation this Article presents, is problematic. Although bibliographic data is used to advance knowledge, and in one sense is knowledge, it is not an expression of "ideas" in the usual sense.\(^2\) Thus, this factor will not be determinative on the question of fair use.

The second factor is the substantiality of the copying. Here the amount of the copyrighted work copied will be very small. In discussing the issue of data base proprietors' rights in the product of searches, CONTU concluded that "the issue of how much is enough to constitute a copyright violation would likely entail analysis on a case-by-case basis with considerations of fair use bearing on whether the unauthorized copying of a limited portion of a data base would be held non-infringing." CONTU did note that "[f]air use should have very limited force when an unauthorized copy of a data base is made for primarily commercial use." It went on to say, however, that only "if information of a substantial amount were extracted and duplicated for redistribution would serious problems exist, raising concerns about the enforcement of proprietary rights."\(^2\)\(^4\) Thus "amount copied" will not in itself be a helpful criterion.

Examining the amount copied is actually one way of testing how much harm has been done to the proprietor's economic interest, that is, the third statutory factor. Furthermore, "[e]conomic damage to the proprietor of a data base is precisely that which would result from an unauthorized search: the amount of information actually retrieved has little if anything to do with recognition of the proprietor's rights."\(^2\)\(^4\) In the example, an unauthorized search is not being posited, but rather distribution of the results of a search which may decrease the market for search services. In this sense, Harvard is directly competing with the search service by providing the same results. Lockheed might therefore be able to show a likelihood of economic harm, but showing actual harmful effect would be difficult, for the number of students or even faculty willing to pay for the search is questionable. Sale to the public would mean greater interference with Lockheed's potential market. As a commercial purpose, it would almost certainly make Harvard an infringer. Pursuant to \textit{Sony}, a commercial purpose creates the presumption that there is a likelihood of economic harm to the compiler. Mere non-

\(^{239}\) See \textit{supra} note 193 and text accompanying notes 218-23.

\(^{240}\) Final Report, \textit{supra} note 95, at 42.

\(^{241}\) Squires, \textit{supra} note 55, at 232.
commercial distribution, or even distribution for a fee covering costs, would present a stronger case for fair use.

Harvard could find further encouragement from Williams & Wilkins v. United States, where the same opposing factors of direct competition by a non-profit institution and problematic economic harm to the publisher were present, as well as from the Supreme Court decision in Sony v. Universal.242 In Williams & Wilkins, the Court of Claims found photocopying of journal articles by the National Library of Medicine and the National Institute of Health to be fair use, stressing the importance of the use for medicine and medical research. The amount of copying was startling: in one year, NIH copied about 93,000 articles; NLM about 120,000.243 The court took a strict view of the showing of economic harm necessary and found for the defendants since the plaintiff had “never made a detailed study of the actual effect of photocopying on its business, nor had it refuted defendant’s figures” (figures that showed the libraries had been increasing their journal subscriptions and that plaintiff was a large, profitable, growing business of which the four journals in question were a very small part).244 As one commentator pointed out, under such a test a finding of fair use would be almost inevitable given the non-profit nature of the use.245

Lockheed would probably be no more able than Williams and Wilkins were to prove actual economic harm from copying of the search results, if Harvard avoided commercial distribution. This points up the fact that “economic harm to the copyright owner” and the commercial nature of the use are often ways of expressing the same concept: for-profit use is much more likely to cause economic harm to the copyright owner by decreasing the potential market.

Harvard’s purpose in merely copying the searches for faculty and students could rather easily be characterized as “advancement of knowledge.” Harvard would then be very likely indeed to prevail in its defense of fair use.

Is this a defensible result? True, the creator of bibliographic

242. Williams & Wilkins, 487 F.2d 1345 (Ct. Cl. 1973), aff'd by an equally divided court, 420 U.S. 376 (1975); Sony, 44 S. Cr. Bull. (CCH), at B641. Although in Sony the court stated that actual harm need not be shown, a plaintiff in a situation such as that in Williams & Wilkins would still bear the burden of showing a “meaningful likelihood” of harm from a library’s copying. More specifically, in challenging a noncommercial use the compiler must show that a particular use is harmful, or that if it should become widespread, it would adversely affect the potential market for the copyrighted work. Id. at B674.

243. 487 F.2d at 1373-74.

244. Id. at 1357-58.

data,\textsuperscript{246} regardless of the skill and judgment used in arriving at the bibliographic description, does not invest the amount of effort that is involved in the typical copyrighted work. Also, free dissemination of bibliographic data is an essential part of the research process. Dissemination thus seems to serve the ultimate purpose of the copyright statute. However, creation of bibliographic data is expensive. "Persons must first read the printed material, then analyze, index and otherwise code it, and put it into machine-readable form."\textsuperscript{247} The commercial creator needs some incentive to undergo this production expense.

It has been shown that courts recognize the need to preserve economic incentives for creation since they focus on the economic harm to the copyright owner that may result from use when they examine whether a fair use privilege exists. If liability turns essentially on the question of economic harm, the law should not differentiate between machine-readable compilations of bibliographic data and hard copy compilations, unless the economics of the data base situation are different.

The question of fair use in connection with hard copy compilations will arise only in those cases where an infringement has been found, that is, where there is substantial copying of the owner's work by the user. A second industrious bibliographer is free to do his own original compilation by gathering bits from many bibliographies and using them in a new work. Assuming that the individual entries taken were not copyrighted, and that no one's compilation was substantially copied, the work would not infringe the rights of anyone, even if sold for profit. Even if it directly competed with an existing bibliography, it might be protected if independent effort beyond mere copying could be shown. The professor or librarian using printed bibliographies to compile a new one for later distribution or even sales would thus be protected from any charge of infringement.

This being so, it seems that if a library pays for searches of several data bases on particular subjects and combines the results of those searches into a new bibliography, there should be no finding of infringement. However, one can argue that if the library uses its own in-house computer to do the searches, and makes its new com-

\textsuperscript{246} "Bibliographic data" comes in several varieties; the key word, author, and title entries provided by commercial indexes are far less appealing subjects for protection than the cataloger's careful description with subject headings and classifications. But many commercial services abstract articles—certainly an exercise as creative as cataloging. Therefore, the commercial nature of the creator will not be useful in evaluating the extent of the creator's right, for it seems that it should not by its commercial nature alone be automatically due less protection.

\textsuperscript{247} O.T.A., \textit{supra} note 165, at 164.
pilation by either merging the search results or changing their format, then it has engaged in no original work and is thus unprotected from liability. The machine-readable nature of the data also decreases the possibility of finding original “arrangement” in its compilation and makes protection rest on a “labor expended” rationale. Although the nature of the labor has changed, the gathering and input of data still involves a significant expenditure of time, labor, and money. Thus, the data base should be protectible by copyright. If the library has expended the necessary labor and not substantially copied any data base, then its use should be protected from a finding of infringement.

Even if infringement were found, the library could argue fair use in that there was no significant economic harm to the publisher. Although the new bibliography might temporarily eliminate the need to search a data base on one particular subject, the number of subjects covered by the data base would make that decrease in demand minimal.

The data base vendor does have one argument that makes his case stronger than that of the hard copy publisher. If a library’s new bibliography eliminates the need to look at a hard copy bibliography in the library’s collection, then the use of the hard copy will decrease, but the publisher will suffer no economic harm, assuming that users would not ordinarily buy a hard copy bibliography just to consult it. Sales to libraries will remain constant. However, if a library’s bibliography eliminates the need to search a data base, then the data base publishers and licensed vendors have lost fees for one search per user. Of course, this perceived difference rests on an assumption that users would pay for the data base searches but not for hard copy bibliographies in the absence of the library’s new bibliography.

If the assumption is accurate, data base publishers could make a stronger showing of economic harm, which might weigh against

248. See supra text accompanying note 218.
249. Oberman, supra note 61, at 13-14: “The basic structure of the directory must be conceived, a computer program developed, an information input procedure arranged, and a final format for the directory designed. . . . There may be less ‘sweat of the brow’ but there are at least compensating administrative headaches to be overcome.”
250. Of course, it might be liable for breach of contract under the standard terms imposed by data base vendors upon users, which restrict such uses. See supra note 155.
251. The idea of hard copy as “free” and data base as “fee” reflects the distinction libraries themselves have tended to make and raises interesting questions about who pays and should pay for library services, questions that are unfortunately outside the scope of this Article.
finding a fair use privilege in the library use of search results from several data bases in a bibliography. However, the question is unlikely to be reached, since the preliminary showing of infringement would be difficult to make where the library has contributed original organization and labor to data culled from a number of sources.

More troubling is the case of "iterative" copying done by the library, as in the hypothetical where the library reproduces search print-outs, thus directly substituting copies for a search of the data base. Here the difference between data base and hard copy seems to vanish. Publishers assume that users would subscribe to journals if libraries refused to make copies of articles or refused to provide copying facilities. One can say that since Williams & Wilkins held library copying of articles on a large scale to be fair use, libraries may likewise copy data base print-outs. That holding, however, rested in good part on plaintiff's inability to show economic harm.\footnote{252}{See Gordon, supra note 245, at 1652, for a criticism of this test.} After Sony, a plaintiff might easily be able to show a meaningful likelihood of future harm from the library's activity, thus barring a finding of fair use.

One alternative test of economic effect, proposed in a recent law review article,\footnote{253}{Id. at 1600.} would require a finding of market failure before allowing fair use. The market fails when transaction costs exceed anticipated benefits, or when benefits from the use are not all translated into compensation to the user, so that the market price he can pay the owner is lower than it should be. The market also fails when use contributes value that is not easily monetized, or where the copyright owner refuses to allow use only because he wants to restrict dissemination of the information.\footnote{254}{Id. at 1627-35.} Gordon would use two additional criteria in a fair use test. \textit{Id.} at 1614. The second step after finding market failure would be to determine whether the use was in the public interest, that is, whether the benefits of defendant's use outweighed injury to plaintiff. Finally, the use by defendant would have to be shown not to interfere substantially with plaintiff's incentive to create. These last two factors seem little different from the traditional fair use considerations discussed previously, and will therefore not be analyzed further here.

\footnote{252}{See Gordon, supra note 245, at 1652, for a criticism of this test.}
\footnote{253}{Id. at 1600.}
\footnote{254}{Id. at 1627-35.}
fact be cured.\textsuperscript{255} Thus, where high transaction costs produce market failure, a centralized clearinghouse (similar to those operated by ASCAP and music publishers) might eliminate the problem. If such a cure is possible, Gordon believes that there is no need for a fair use privilege. Therefore, she criticizes the court in \textit{Williams & Wilkins} for not considering the publishers' suggestion of a clearinghouse as a solution, though she recognizes the difficulties inherent in a court's imposing such a scheme.\textsuperscript{256}

In the hypothetical situation, the factor of research value might outweigh the publishers' suggestion, as it did in \textit{Williams & Wilkins}. Where the economic harm is not very clearly shown, and the use is non-commercial and part of the process of advancement of knowledge, it does seem that fair use is a justifiable result, even under an economic test such as market failure. The fair use doctrine will not protect the commercial creator from non-commercial uses such as the library's use in the hypothetical.

The data base proprietor does have another avenue of protection—a pragmatic one. Since he cannot monitor use of the search print-out by the searching institution, he can take account of the possible harm to his data base market from that use in the prices he charges for the search itself. Use of the data base is easily controlled and monitored.\textsuperscript{257} In fact, pricing differentials between institutions and individuals are the rule in the case of academic journals, and the economic calculations involved have been applied to data bases.\textsuperscript{258}

4. \textbf{Compulsory Licensing As a Possible Alternative}

If a data base proprietor can protect itself from economic harm by pricing its services accordingly, why is traditional copyright protection (with exclusive right to reproduce in the owner) needed? Compulsory licensing is an alternative that has been considered in connection with computer-related works.\textsuperscript{259} The concept of compulsory licensing was established under the 1909 Act for recorded musical compositions to prevent absolute control of popular music by its copyright owner. Under compulsory licensing, whenever the copyright owner authorized public distribution of a record embodying that composition, any other person could make a record embodying

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{255} Id. at 1620-21.
\item \textsuperscript{256} Id. at 1623.
\item \textsuperscript{257} R. Saltman, \textit{supra} note 161, at 34.
\item \textsuperscript{258} Braunstein, \textit{The Role of Copyright Protection and Optimal Pricing in Computerized STI Systems}, app. D in R. Saltman, \textit{supra} note 161.
\item \textsuperscript{259} See, e.g., R. Saltman, \textit{supra} note 161, at 58, quoting in part CRC Systems Inc., \textit{supra} note 120, at 190, 194-204.
\end{itemize}
\end{footnotesize}
the composition provided that he paid a statutory fee for each copy. In the 1976 Copyright Act, the concept was extended to cable system transmissions of broadcast programs, non-commercial broadcasters' transmissions of certain works, and jukebox performances of music.\textsuperscript{260}

The compulsory license seems particularly well suited to bibliographic data. The Network Advisory Committee, in discussing the issue of ownership of bibliographic data, noted that libraries are not very interested in "ownership" of individual cataloging records. Rather, they are concerned with gaining access to a source of records.\textsuperscript{261} Under compulsory licensing, the creating library could not prevent access to the bibliographic data, but the user would have to pay for that use. Compulsory licensing would produce the same result as the optimal private agreement in the Harvard-Yale situation discussed above: Harvard could use Yale's record, but would have to pay for its use. Likewise, in the second situation, RLIN could not prevent the compiler's use of tapes from its data base, but he would have to pay for them. Commercial vendors already allow anyone to access their data bases upon payment of a fee, but they are not "compelled" to allow any particular use. Presumably compulsory licensing for them would mean raising fees and allowing broader uses of search results—not a very different result from that presently available under private contractual arrangements.\textsuperscript{262}

However, compulsory licensing is not a panacea. One commentator notes that because it constitutes a loss to the holder of the copyright it could conceivably decrease his incentive to create and thus thwart the purpose of copyright.\textsuperscript{263} "Further, an objectively determined 'price' is always less satisfactory to a property owner than a price of his own choosing and is usually a less reliable indicator of value for society's purposes as well."\textsuperscript{264}

Moreover, compulsory licensing introduces an element of regulation with resulting complexity and transaction costs which may be undesirable.\textsuperscript{265} While use of a data base and consequent license fees may be easy to keep track of, there will be need for administra-


\textsuperscript{261} D. Webster & L. Maruyama, supra note 83, at 15, 17.

\textsuperscript{262} See supra text accompanying note 155.

\textsuperscript{263} Y. Braunstein, supra note 162, at 243.

\textsuperscript{264} Gordon, supra note 245, at 1623.

\textsuperscript{265} See R. Saltman, supra note 161, at 34-36. Royalty levels are established by statute, with the Copyright Royalty Tribunal having authority to change the rate structure within limits. See D. Johnston, supra note 260.
tion, reporting, and possible arbitration of disputes by a government agency. "From the point of view of the copyright owners . . . a system that permitted certain limited uncompensated takings to occur . . . might be preferable to a system in which compensation was guaranteed but only after the fact."266 In some cases, a library or a non-profit data base owner might wish to prevent particular users from gaining access to its bibliographic data.267 Compulsory licensing deprives the creator of any control over the destiny of his work.268

Finally, application of fair use to a compulsory licensing scheme would be problematic. The whole idea of the compulsory license is that no distinctions among users should be allowed. Exemptions for some would mean that others would bear the burden of higher fees.269 Thus, even if the compiler in the second hypothetical situation were a non-profit organization, it would arguably have to pay the license fee for use of the RLIN data base. On the other hand, fair use is a limit on any owner's rights—in a sense it goes beyond compulsory licensing and allows "compulsory free use." If so, one is faced with the problem of economic disincentive that fair use presents in connection with statutory ownership rights.

One answer to this is that the rigidity of compulsory licensing may be avoided by private contract just as the uncertainty of copyright law now is.270 However, compulsory licensing with its ready-made price structure might discourage private arrangements and would certainly make it impossible for a creator to gain more than the statutory license price regardless of circumstances meriting increased compensation.271

5. Appropriateness of Fair Use As A Mediating Doctrine: Conclusion

Like compulsory licensing, existing copyright law for computer data bases provides a background with known criteria which parties

266. Gordon, supra note 245, at 1623.

267. For example, a library which creates a large percentage of original records might not wish to allow a "freeloading" library that creates none to have access to its records. See supra note 204.

268. Y. BRAUNSTEIN, supra note 162, at 243.

269. Id. at 243-44.

270. The terms of a compulsory license may be waived by contract, at least in the case of phonograph records and presumably in general. See M. NIMMER, supra note 26, at § 8.04[1].

271. See supra text accompanying note 264.
may take into account in reaching their own agreements. Unlike compulsory licensing, however, its criteria are flexible and may be variously weighed in each individual set of circumstances. The fact that the fair use doctrine allows a court to reach opposite results in the case of the compiler above makes planning difficult and increases uncertainty costs. Conversely, it reflects the fact that different results may be appropriate even in quite similar circumstances. The public interest in access to information may after all be served in some situations by granting copyright with its incentive to produce, in others by refusing copyright as too great a bar to dissemination and new production.

For library-created bibliographic data, flexibility in applicable copyright law is most appropriate and necessary. Libraries, concerned with information as a public good which should be as widely disseminated as possible, have traditionally desired to share the bibliographic data which helps provide access to that knowledge. The changing technology of cataloging, reflecting the potential for use of bibliographic data as a commodity in itself, has caused libraries to become concerned about rights in bibliographic data they create with respect to other libraries and non-libraries who would use it. Nevertheless, the library's interest in bibliographic data does not fit into the standard categories of copyright ownership. Libraries and the bibliographic utilities have thus worked out a variety of arrangements balancing the costs and benefits libraries incur in creating and receiving bibliographic data.

Agreements aside, without the mediating doctrine of fair use, the law of copyright ownership produces untoward results when applied to library situations. With the doctrine of fair use, and with freedom from the rigidity of compulsory license requirements, copyright law can balance the competing interests of creator and user in each situation, and thus ultimately achieve the promotion of knowledge and the growth of wisdom in the world.