Databases and Dynamism

Michal Shur-Ofry
Hebrew University of Jerusalem

Follow this and additional works at: http://repository.law.umich.edu/mjlr
Part of the Computer Law Commons, Intellectual Property Law Commons, and the Internet Law Commons

Recommended Citation
Available at: http://repository.law.umich.edu/mjr/vol44/iss2/2

This Article is brought to you for free and open access by the University of Michigan Journal of Law Reform at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in University of Michigan Journal of Law Reform by an authorized editor of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.
DATABASES AND DYNAMISM

Michal Shur-Ofry*

Databases are generally perceived in legal scholarship as static warehouses, storing up valuable facts and information. Accordingly, scholarship on copyright protection of databases typically concentrates on the social need to access their content. This Article seeks to shift the focus of the debate, arguing that the copyright-databases debate is not merely a static "access to information" story. Instead, it is a dynamic story of relations, hierarchies, and interactions between pieces of information, determined by database creators. It is also a story of patterns, categories, selections, and taxonomies that are often invisible to the naked eye, but that influence our perceptions of the world in manners of which we are seldom aware.

Relying on socio-psychological literature and communication theories concerning complexity, categorization, and stereotyping, this Article examines the dynamic dimension of databases. It argues that this narrative should direct legal attention toward the protection afforded by copyright not to contents of databases, but rather to their "selection and arrangement"—an element which has been largely ignored by legal scholarship. While the Article does not advocate a complete expiry of copyright in "selections and arrangements," it does hope to spark a discussion with respect to their social and economic role, and add a new dimension to the copyright-database debate.

INTRODUCTION

Think of your favorite legal database—the one which is your first choice when searching for legal materials. Now, imagine that you have to perform an urgent search, but for some reason cannot access that database. You have to use another database, which contains identical content, but which is arranged in an unfamiliar way. The frustration, waste-of-time, and helplessness that you are likely to experience while trying to figure out the structure of that other database are at the center of this Article.

During the last two decades, copyright law has been engaged in a heated, cross-national debate over the protection of databases. Sparked by the Supreme Court's landmark decision in Feist

---

* Hebrew University of Jerusalem. For valuable comments and helpful discussions, I thank Katya Assaf, Dan Burk, Daphna Lewinsohn-Zamir, Guy Pessach, Steven Wilf, the participants of the WIPO Seminar on Intellectual Property and Creative SMEs in the Digital Environment (Geneva, 2009), and the editors of this Journal. The opinions expressed here are solely my own.
publications, Inc. v. Rural Telephone Service Co., this discussion has largely focused on the contents of factual compilations. It has been commonly phrased as an "access to information" debate, and has been associated with famous catchphrases such as "information wants to be free."2

Whether the positions expressed by scholars and policy makers are for or against Feist, the general narrative emerging out of that discussion draws databases in a similar manner: big, static warehouses, storing facts and information that their owners seek to lock up, relying on copyright law. Feist, which held that factual contents of compilations do not enjoy copyright protection regardless of the investment—or "sweat of the brow"—expended in their assembly, is hence perceived by its many supporters as an unlocking mechanism: the decision provides access to important facts and information, while offering database owners the comfort of copyright protection in their "selection and arrangement," if the latter possess a "minimal degree of creativity."4

Yet, databases are not merely static piles of facts and data. By selecting the information which enters the database, and organizing that information in a recurring and systematic manner, database creators determine both what information users of a database can access, and also how this information will be communicated to the user. The story of copyright and databases, then, is not merely an

3. For a detailed discussion of the "sweat of the brow" doctrine which has developed by pre-Feist courts, see Feist, 499 U.S. at 352 ("[T]he underlying notion was that copyright was a reward for the hard work that went into compiling facts . . .").
4. Id. at 362.
access to information story. Rather, it is a story of dynamic structures and patterns that determine relations and define hierarchies between pieces of information—a story of organization, taxonomy, and categorization of information that influence our perceptions of the world in manners of which we are seldom aware. Indeed, even the driest and dullest database, such as a yellow pages directory, is a site of social dialogues and power-relations. Consider, for example, an entry in a yellow pages directory providing information about “Eating Disorders Centers,” which can be catalogued under the broader category of “rehabilitation,” “treatment,” or even “emergency.” Such a categorization does not merely provide access to a particular piece of information. Rather, it embodies ineluctable assumptions regarding the significance and social context of that information, which can influence user comprehension and shape perceptions in a broader sense.\(^5\)

This Article focuses on this overlooked narrative, using sociopsychological insights and communication theories which highlight the cognitive and communicative significance of categorization, selection and classification. This literature clarifies the role of familiar patterns and formats as complexity reduction mechanisms; it explains human dependency upon such structures and schemas and suggests that the relations, hierarchies, and stereotypes created by the systematic classification of information have strong cognitive and social impact. Unsurprisingly, then, structures and arrangements of databases are susceptible to economic phenomena of standardization and user lock-in.

Highlighting this overlooked narrative shifts the focus of the copyright-databases debate, from control over contents to control over selection and arrangement. While the issue of copyright in factual contents of databases has been fiercely debated in legal literature, broad copyright protection of database selections and arrangements is taken almost for granted, generally deemed desirable, and has hardly been explored in copyright scholarship.\(^6\) This Article seeks to fill the gap. It reveals the role of selections and arrangements as powerful paratexts,\(^7\) which intermediate between

---


7. The term “paratext” was recently introduced into copyright’s discourse by Dan Burk. See Dan L. Burk, Copyright and Paratext in Computer Gaming, in Emerging Ethical
the user and the content of the database, and influence the manner in which that content is perceived and interpreted. It suggests that the difficulties entailed in affording broad copyright protection to database selections and arrangements can be substantial and influential, and should not be ignored. It further argues that overlooking the issue of copyright in selection and arrangement not only burdens the handling of those difficulties, but also yields incoherencies within copyright doctrine.

By focusing on the dynamic narrative of copyright and databases, this Article advocates a more cautious approach toward broad copyright protection of selections and arrangements. It further hopes to spark a discussion which will shift—at least in part—the focal point of the debate from content to selection and arrangement.

The Article proceeds as follows: Part I uses socio-psychological scholarship and communication theories to explore the social role of structures and categories in general, and of database selections and arrangements in particular. It continues to examine the economic implications of this analysis, and highlights the inclination of database selections and arrangements toward standardization. It then examines the costs and externalities associated with this dynamic dimension from both normative-cognitive and economic perspectives. Part II then builds on these insights to analyze the difficulties caused by copyright law's current disregard of this dimension and to examine possible implications of the analysis on the scope of copyright protection afforded to database selections and arrangements. Concluding remarks follow.

I. DATABASES' DYNAMIC DIMENSION

Databases are not merely pieces of information and data piled together. Rather, databases offer their users a systematic organization of certain information. The creation of a database involves a process of choice, classification, and categorization of information, or—to use copyright terminology—a process of selection and ar-

---

ISSUES OF LIFE IN VIRTUAL WORLDS 33 (Charles Wankel & Sean Malleck eds., 2010). The term is attributed to Gérard Genette, who uses it to describe ancillary texts and devices that mediate between the reader and the primary text. Gérard Genette, Paratexts: Thresholds of Interpretation (Cambridge Univ. Press 1997) (1987).

8. See Council Directive 96/9, art. 1.2, 1996 O.J. (L 77) 20 (EC) (defining a “database” in the European Union Database Directive as “a collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means” (emphasis added)); see also Bowker & Star, supra note 5, at 10 (pointing out that consistency is an important attribute of classification systems).
Databases and Dynamism

This process is comprised of a series of decisions and judgments regarding the scope and nature of material to be indexed, the categories to be included, and the items to be assigned to each category. In the pre-digitized era, the structure and scope of most databases were often apparent to their users. Nowadays, when many important databases are electronic and on-line, such structures are seldom visible to the ordinary user. Yet, as demonstrated below, in digital databases too, the selection and arrangement of information is necessary, unavoidable, and no less influential.

The process of selection and arrangement determines not only the set of rules according to which items are stored in and retrieved from the database, but also the relations, links, and hierarchies between those items. By so doing, it contextualizes the database's content and facilitates the comprehension and processing of that content by the database's users. In the following sections, I refer to these features as the dynamic-interactive dimension of databases.

A. On Categorization and Standardization

Socio-psychological research and communication theory both indicate that the selection and arrangement of database content facilitates the comprehension and processing of that content by users. Scholarship in these fields describes categorization as "the process of understanding what something is by knowing what other things it is equivalent to and what other things it is different from." This literature further suggests that the process of selecting, indexing, and categorizing data is crucial to our ability to process, communicate, and understand information.

9. See Bowker & Star, supra note 5, at 33-35 (arguing that usable classifications are invisible "almost by definition"); see also discussion infra Part I.  
10. See discussion infra Part I.B.  
12. See infra notes 13-14 and accompanying text.  
13. Bowker & Star, supra note 5, at 290 ("We know what something is by contrast with what it is not."); see also Craig McGarty, Categorization in Social Psychology 1 (1999).  
In semiotic terms, the relationship between database structure ("selection and arrangement") and content is analogous to the relationship between *langue* and *parole*. The predetermined set of rules and conventions expressed by a database's selection and arrangement forms the database's *langue* (language), while the specific contents and data assigned to each category are particular instances of *parole*, or "speech," in that language. Thus, much like spoken languages, our ability to comprehend a specific *parole* depends upon our familiarity with the *langue*, or in our case, with the syntax and structure of the database. Users become exposed to the database's *langue*—or selection and arrangement—during their search for specific data or information. Consider online legal databases: users of Westlaw International searching for legal materials from New Zealand would learn to find them under the category of "Asia and the Pacific Rim," while users of Lexis-Nexis would find New Zealand under the "European Union, Commonwealth and Foreign Nations" section. This exposure to the database format is often inadvertent, as the attention of the user is normally focused upon the materials sought. Yet over time and repeated uses, the *langue* of the database becomes familiar and is imprinted upon its users' minds. At that stage, users become dependent upon their database's selection and arrangement.

These latter insights are strengthened by research in cognitive and social psychology, which indicates that reliance upon familiar patterns and categories is a vital cognitive tool. This literature regards categories as complexity reduction tools that allow us to quickly identify individual and related information and to process the large (and otherwise overwhelming) amounts of data in our complex environment. Familiar categories and schemas context-

---

15. For the semiotic distinction between *langue* and *parole*, see Daniel Chandler, Semiotics: The Basics 12-14 (2002); Ferdinand de Saussure, Course in General Linguistics 11-17 (Charles Bally & Albert Sechehaye eds., Wade Baskin trans., 1959).
17. Follow this path in the Lexis-Nexis Academic database: Legal International > EU, Commonwealth, and Other Nations > New Zealand Cases.
19. Allport, supra note 14, at 20; McGarty, supra note 13, at 25, 29, 45.
Databases and Dynamism

tualize the information to which we are exposed, and facilitate our processing and comprehension of that information. Consider for a moment the following four sentences taken from a well-known psychological experiment:

The procedure is actually quite simple. First you arrange things into different groups depending on their makeup. Of course, one pile may be sufficient depending on how much there is to do. If you have to go somewhere else due to lack of facilities that is the next step, otherwise you are pretty well set.

Most likely, this information would be much more comprehensible if placed in a context—in this scenario, by adding the title “Washing Clothes.” This example illustrates that categories and schemas formed in our minds are necessary for perception: their use enables us to form expectations and knowledge, and makes our world a “reasonably predictable place.”

In light of their cognitive function, familiar categories are impressed upon our minds. Research even indicates that we tend to preserve patterns and schemas after our memory of the precise material has long disappeared: the influential psychologist and scholar Frederic Bartlett maintained that particular experiences are stored in organized settings, and that memories are constructed and reconstructed on the basis of such schema. Other cognitive research demonstrates that categorization is closely related to our ability to memorize and retrieve specific information.

20. See supra notes 13-14 and accompanying text.
22. Id. at 722-25 (demonstrating that participants exposed to the contextualized information, i.e., the text together with the title “washing clothes,” presented increased levels of comprehension and recalling of the text, in relation to participants exposed to the text alone).
23. See Eysenck, supra note 14, at 213-14; see also Allport, supra note 14, at 20-21 (making a similar observation).
25. See generally Geoffrey R. Loftus & Elizabeth F. Loftus, Human Memory: The Processing of Information 66-69 (1976) (presenting research indicating that organizing and structuring information increases the amount of information that people can recall). For an interesting example of the relations between memory and categorization, see Harlene Hayne, Carolyn Rovee-Collier & Eve E. Perris, Categorization and Memory Retrieval by Three-Month-Olds, 58 CHILD DEV. 750 (1987) (demonstrating that categories facilitate
These insights seem directly applicable to categories that are formed and expressed by the selection and arrangement of familiar databases. The repeated use of a database's structure and the familiarity with its \textit{langue} enables users to concentrate on the specific information sought (\textit{parole}) and more quickly identify the required data.\footnote{These similar observations appear in semiotic literature discussing the distinction between \textit{langue} and \textit{parole}. See \textit{Chandler}, supra note 15, at 12-15; cf. \textit{Alastair Fowler, Kinds of Literature: An Introduction to the Theory of Genres and Modes} 20-22 (1982) (making a similar point with respect to literary patterns).} At a certain point, the database's structure and organizational choices diffuse into the information system: for the acquainted user it becomes "natural" and transparent.\footnote{Id. at 323, 325.} We no longer pay attention to the \textit{langue} of the database. We take it for granted. To borrow the description used by Geoffrey Bowker and Susan Star, the database's selection and arrangement becomes invisible.\footnote{For an analysis of the network-communicative attributes of copyright protected cultural patterns, see Michal Shur-Ofry, \textit{The (Copyright) Law of Genre: A Network Perspective on Copyright Protection of Cultural Genres}, 2 FLA. ENT. L. REV. 60, 75 (2008) (discussing the network and communicative attributes of cultural formats). See also Samuelson, supra note 6, at 195-96, 218 (pointing out the communicative aspects of coding systems). The similarities between databases' "selections and arrangements" and other copyright protected patterns and formats are discussed in Part II.B infra.}

This analysis clarifies that, like other types of popular patterns and formats, familiar selections and arrangements have a strong interactive value: the well-known structure serves as an interactive tool that eases the communication between the database and its users.\footnote{See \textit{McGarty}, supra note 13, at 31-45 (discussing various links between categorization and information retrieval).} Consider again your favorite legal database; it is certainly much easier to concentrate on the object of your search (New Zealand materials, for example), if you are already familiar with the arrangement of the database ("New Zealand" materials under "European Union, Commonwealth and Foreign Nations"). Searching identical materials using a different database whose structure is still unfamiliar is likely to prove more difficult and frustrating.

Familiar databases facilitate interaction not only between the user and the database, but also among various users of the same database. Therefore, the use of a familiar structure is significant in a strict economic sense, too.\footnote{In light of the communicative attribu-}
utes and cognitive benefits of categories and classifications, database selections and arrangements are susceptible to economic network effects: widespread use of a database allows people to use it from different computers in various locations, reducing learning costs and increasing the utility of the database for each individual user. In addition, the uniform selection and arrangement allows users to interact with other users. The European IMS case clearly illustrates this latter point: IMS Health, a leading supplier of pharmaceutical market information, used a database for pharmaceutical data reports, which divided the territory of Germany into 1,860 geographical "bricks." The structure became a standard among pharmacies and hospitals in the German market. The use of a single arrangement for pharmaceutical reporting facilitated communication between pharmacies and IMS, which processed the data, and reduced learning costs for the reporting pharmacists.

Similar network advantages (albeit on a lesser scale) are present when users refer each other to information contained in a certain database, such as a restaurant in a restaurant directory, a television program in a TV guide, or a specific address in a street-guide. When both individuals use the same database, the requested information will appear in a similar manner (on the same web-page or hard-copy page, under the same category), and reference will be simpler. Notably, in all these cases, the network advantage lies in the databases' selection and arrangement, and not in their content. A database with identical content but differently arranged would not provide its users with equal benefits.

The cognitive and economic advantages derived from repeated and widespread use of uniform selections and arrangements shed light on the inclination toward standardization in the database

31. For the economic concept of network effects in general, see Michael L. Katz & Carl Shapiro, Network Externalities, Competition, and Compatibility, 75 AM. ECON. REV. 424 (1985).
33. Id. at I-5072.
34. Id.; Koelman, supra note 6, at 828.
35. Koelman, supra note 6, at 828. For the network value of other pharmaceutical classifications, see BOWKER & STAR, supra note 5, at ch. 3 (discussing the ICD classification of disease and related health conditions designed by the World Health Organization).
industry. As demonstrated above, the use of a standard selection and arrangement facilitates interaction and allows database users to benefit from network advantages. This analysis may explain the dominance of a mere few players in the legal databases industry, even though the content in these databases (legal materials) is often free of copyright protection. It also sheds light on the standard arrangement of legal materials which emerged in that industry. It similarly explains IMS's standard status in the German pharmaceutical information market, as well as the dominance of a Melbourne street guide, which was the subject of the Australian Melway Publishing Pty. Ltd. v. Robert Hicks Pty. Ltd. case. Likewise, the continued popularity of the Ticketmaster database in the online ticket-selling industry, despite court decisions which have repeatedly held that its content is not copyright-protected and may be freely extracted by competitors, may also be attributed to user dependence upon that database's selection and arrangement.

Yet, in spite of the crucial cognitive and economic role just described, the dynamic-interactive dimension of databases is not free of social costs. The following paragraphs outline the impacts and externalities associated with this dimension. Part II then continues to examine the role of copyright law in this context.

36. See Carl Shapiro & Hal R. Varian, Information Rules: A Strategic Guide to the Network Economy 122–23 (1999) (discussing standardization in the database industry and its relation to network effects); cf. Bowker & Star, supra note 5, at 15 (pointing out the connection between classifications and standards); Koelman, supra note 6, at 823–25 (discussing network effects leading to standardization in utilitarian standards, including in the context of the IMS database). This inclination towards standardization is, of course, closely related to the "lock-in" phenomenon. See infra Part I.C.

37. See Matthew Bender & Co. v. West Publ’g Co., 158 F.3d 674, 683–85 (2d Cir. 1998) (noting the existence of an industry standard in the presentation and arrangement of legal cases, consisting, inter alia, of standardized format for party names, restyling of court name and relevant dates and presentation of caption, court, docket number, and date in a particular order).

38. (2001) 178 A.L.R. 253 (concerning a dispute between the publisher of a Melbourne street directory which constituted a de-facto standard, and a former distributor that required continued access to the distribution of that directory); see also Richardson, supra note 31 (highlighting the network effects subsisting with respect to the database which was the subject matter of the litigation).

B. Databases as Normative Frameworks

The discussion above demonstrates that the dynamic-interactive dimension of databases stems from our human need to categorize and contextualize information. It also implies that databases have a meaning-making function, and a normative dimension. Cognitive research clarifies that defining different classes and allocating information into them enable people to generate predictions and expectations about the members and components of those categories. By so doing, classes and categories create normative frameworks that influence people’s expectations, affect the ways in which material is interpreted and comprehended, and shape our perception of reality.

Notably, the meaning-making site is not the mere facts and data included in a database. Rather, it is the selections and arrangements that serve as paratexts and intermediate the database content to its users. By determining the scope of information included in the database and the manner in which that information is accessed and retrieved, selections and arrangements contextualize database content and influence the manner in which that content is understood and interpreted by users. As noted by Niva Elkin-Koren, information structures impose upon their users the judgment of their creators regarding the meaning and importance of that information. Consider the listing of “Alcoholic Anonymous” in a yellow pages directory: listing this item in the “Emergency Services” section is a different meaning-making act than placing this very item in another section such as “Rehabilitation.”

However, this impact of selections and arrangements is not confined to physical databases. It remains valid with respect to computerized online databases that govern our digital lives. Digital databases offer users more flexible search options, including “free

40. See supra note 14 and accompanying text.
42. Eysenck, supra note 14, at 215–18; McGary, supra note 13, at 254–55, 261–62; see also Bowker & Star, supra note 5, at 46–49 (pointing out that classifications become authoritative in describing the universe).
43. Cf. Burk, supra note 7, at 17–18 (making a similar observation with respect to rules of online games).
45. Bowker & Star, supra note 5, at 56–57 (further arguing that current tendency to list Alcoholics Anonymous under “Emergency Services” reflects both the recognition of its reliability as well as the social acceptance of alcoholism as a medical condition). For a similar example, see supra note 5 and the accompanying text.
text" searches. Yet, in light of the overwhelming amount of information covered by such databases, selection and arrangement is inevitable, and its effect remains significant. Consider the Social Science Research Network's recent decision to add a new category to its publications, namely the "SSRN network of elder law studies." This new category does not merely provide an improved search mechanism for those interested in legal issues concerning the elderly. It also embodies an implied statement as to the social importance of this subject, which may have indirect, yet significant, effect upon SSRN's subscribers. Likewise, consider again the different categorization of New Zealand in electronic legal databases. The fact that New Zealand is part of "Asia and the Pacific Rim" according to the organization of one database, while according to the other it is part of the "European Union, Commonwealth and Foreign Nations," does not merely affect the efficiency of searching for legal materials. Rather, these classifications may influence user perceptions in a subtle, yet much broader, sense.

Another prominent example concerns search engine databases. At first sight, it may seem that internet search engines offer the most flexible, free-text-based search, which does not require the user to follow any underlying set of assumptions. But this impression is very much illusory. As Eric Goldman has observed, search engines, too, exercise unavoidable editorial control over their databases. Google, for example, arranges search results by a set of assumptions reflected in its ranking algorithm. Despite the search's flexibility, then, the organization of the information is neither "neutral" nor objective. Rather, it inevitably imposes on its users the underlying judgment of its creators. The assumptions underlying the Google database, for example, assign much importance to popularity in determining a site's ranking among search results.

46. Cf. Elkin-Koren, supra note 44, at 238-40 (arguing that digital online databases are less vulnerable to the control of their creators).

47. Elder Law Studies ejournal, SOCIAL SCIENCE RESEARCH NETWORK, http://www.ssrn.com/update/lsn/lsn_elder-law.html (last visited Oct. 11, 2010). SSRN notes that establishing "elder law" as a distinct category is not an obvious decision. Id. ("Unlike many other areas of the law, elder law is defined primarily by the client population to be served, not by a distinct set of legal doctrines.").

48. See supra notes 16–17 and accompanying text.


The fact that this set of assumptions is not always transparent to the ordinary user does not make it less influential: the popularity-based arrangement of search results increases user exposure to already popular websites, and may strengthen and reinforce majority views among those users.\textsuperscript{51}

Notably, the impact just described is not confined to the arrangement of databases. The selection of the material included in (or excluded from) a database can be equally important in shaping people's views and priorities. Popular databases form prisms through which certain universes of data and information are viewed and analyzed. Their selection may actually come to define the scope of those universes for their users. Many of those searching the internet, for example, are likely never to encounter web pages which are excluded from the databases of prominent search engines.\textsuperscript{52} Likewise, the exclusion of certain materials from a legal database not only reflects the compiler's judgment as to the importance or relevance of such data; rather, these decisions shape the users' legal universe. As illustrated by Daniel Dabney, they may also have a deeper impact upon the development of law itself.\textsuperscript{53}

The normative dimension of database selections and arrangements may be even more significant in light of our general ignorance of it. As noted earlier, database structures tend to disappear with time and repeated use.\textsuperscript{54} With a specific utilitarian task in mind (such as finding a restaurant, legal case, or phone number) we seldom notice the judgments, interpretations, and hierarchies embodied in a database's structure and scope. This influence is further enhanced in light of the database industry's inclination towards standardization. When a database is an industry standard, the repeated exposure of users to its selection and arrangement, often accompanied by a lack of exposure to alternative selections and arrangements, is likely to have an increased impact. The

---


\textsuperscript{52} See id. at 171 (remarking that "to exist is to be indexed by a search engine"); see also Goldman, supra note 49, at 190, 197; cf. Guy Pessach, \textit{[Networked] Memory Institutions: Social Remembering, Privatization and Its Discontents}, 26 Cardozo Arts & Ent. L.J. 71 (2008) (arguing that the digitized networked environment is increasing the influence of private entities on the formation of beliefs and ideas and on our "social remembering").

\textsuperscript{53} See Daniel Dabney, \textit{The Universe of Thinkable Thoughts: Literary Warrant and West's Key Number System}, 99 Law Libr. J. 229, 243-46 (2007) (demonstrating how the scope of material indexed by the West Key Number System, which excludes legal scholarship, may be navigating the development of law in conservative directions).

\textsuperscript{54} See supra notes 26-28 and accompanying text.
judgments and values reflected in those database structures may be perceived as "natural," unquestionable truths. To borrow Dabney's illustration, such popular selections and arrangements shape our "universe of thinkable thoughts."

Databases, then, play a notable—if rather concealed—role in influencing people's perceptions, priorities, and understanding. This role is performed by their selections and arrangements, which function as meaning-making sites and have significant dialogic and democratic implications.

C. Selections, Arrangements, and Lock-Ins

The attributes just described have economic implications too. Standardization stemming from network effects is often characterized by economic lock-in. The ability of an individual user to replace a standard database in a certain market with a new and potentially preferable one depends not only upon the perceived inherent value of both products, but also upon the costs of switching to the new product. These costs are significantly influenced by the magnitude of network effects in the relevant market. In the case of databases, such switching costs include the learning costs which stem from the cognitive dependence upon the familiar selection and arrangement and which can hinder the adjustment to a different structure. They also include costs associated with the inability to interact with other individuals who continue to use the standard selection and arrangement. In order to avoid these costs,

55. See Bowker & Star, supra note 5, at 33, 49 (noting that usable classifications "disappear almost by definition"—"[t]he easier they are to use, the harder they are to see"—and highlighting the tendency to perceive the classifier's description of reality as "true").

56. Dabney, supra note 53 (referring to the effect of the West Key Number System); see also Berring, supra note 50, at 311 (attributing the phrase to Dabney).

57. The latter statement is supported by a large body of literature concerning deliberative democratic theory. According to this model, political will formation is affected by many social interactions, such as reading, deliberation, and interpretation (including, for our purpose, searching and processing of information). The scope of this Article does not allow a comprehensive review of that literature. Prominent writing includes Jurgen Habermas, Three Normative Models of Democracy, in DEMOCRACY AND DIFFERENCE 21 (Seyla Benhabib ed. 1996); Seyla Benhabib, Models of Public Space: Hannah Arendt, the Liberal Tradition and Jurgen Habermas, in HABERMAS AND THE PUBLIC SPHERE 73 (Craig Calhoun ed., 1992); Nancy Fraser, Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy, in HABERMAS AND THE PUBLIC SPHERE 109 (Craig Calhoun ed., 1999); cf sources cited infra note 79 and accompanying text (generally highlighting the dialogic significance of copyright-protected works).

people may continue to use a database which does not necessarily possess the greatest utility and inherent value for them. In other words, users are "locked" in an inferior, yet prevailing, database.

It should be stressed again that declaring the factual contents of a prominent database free of copyright protection does not solve the difficulty just described. The subject matter of lock-in—the component which users become dependent upon and have difficulties extricating from—is a database's selection and arrangement, not its content. Thus, lock-in can occur even when the content of a database is in the public domain or when such content is not compiled by the creator of the database, but added and filled-in by users.

Indeed, the most famous example of lock-in appearing in economic literature, the QWERTY keyboard, can be viewed as an arrangement type of lock-in. According to the prevailing account, the QWERTY keyboard was originally designed during the second half of the nineteenth century to slow down the pace of typewriting, since rapid typewriting during that period resulted in stuck keys on old typewriting machines. Due to significant network effects, QWERTY became the standard throughout the western world. During the 1930's, Dr. August Dvorak introduced an alternative keyboard arrangement. Designed to speed up typing pace, the DVORAK keyboard had been known to be more efficient and user-friendly than QWERTY. Despite these inherent advantages, DVORAK failed to take over the keyboard market. Many scholars attribute this to user lock-in. If DVORAK had been introduced simultaneously with QWERTY, users may well have preferred it due to its inherent value. Yet, once QWERTY has become the standard, the majority of users have decided to cling to it, seeking to avoid the learning costs entailed in switching to an unfamiliar keyboard, and to preserve the advantage of working on identical keyboards from different computers and typewriting machines. The DVORAK keyboard has been practically forgotten (or, to be more

59. See infra notes 72–76 and accompanying text. In many jurisdictions, including the United States, legislation and judicial decisions are not protected by copyright. See 17 U.S.C. § 105 (2006); Banks v. Manchester, 128 U.S. 244 (1888).

60. Such were the circumstances in the IMS case: part of the data "filled-in" in the IMS format is initially supplied by pharmacies and hospitals using the IMS database. See supra notes 32–35 and accompanying text.

61. For a discussion of the QWERTY keyboard, see Geoffrey C. Bowker, Memory Practice in the Sciences 112 (2005); Everett M. Rogers, Diffusion of Innovations 8–10 (5th ed. 2003); Shapiro & Varian, supra note 56, at 185–86; Farrell, supra note 58, at 37–38.

62. Rogers, supra note 61.

63. Id.

64. Id.
accurate, it has been immortalized in network-literature). A quick look at your keyboard will reveal that PC users are still locked in the QWERTY arrangement.

The QWERTY example illustrates that where network externalities are concerned, arrangement of information can be more significant than content: the keys on a QWERTY keyboard cannot be much different than keys on any other keyboard. It is the particular arrangement of that data which has yielded cognitive and economic dependence, and resulted in user lock-in.

II. THE ROLE OF COPYRIGHT

A. Copyright's Impact

Against this background, the significance of copyright protection of "selections and arrangements" becomes apparent. Copyright grants its owners considerable control over a database's structure, provided that structure portrays a "minimal degree of creativity." If that control is too broad, it is likely to increase the social and economic externalities described above.

First, overly broad copyright protection can contribute to economic lock-in and limit users' ability to choose a desired database. When a database is an industry standard, creators of alternative databases wishing to enter the market and offer better products may be unable to do so unless they make some use of the selection and arrangement of the prevailing standard, in order to minimize user switching costs. Such use, however, may constitute copyright infringement.

---

65. *Id.* Notably, the prevailing description in economic literature is strongly disputed by Leibowitz and Margolis, who argue that QWERTY was in fact not inferior to DVORAK and dispute the credibility of other elements in the story. See S.J. Leibowitz & Steven E. Margolis, *Network Externalities: An Uncommon Tragedy*, 8 J. ECON. PERSP. 133, 147 (1994) (arguing that "almost every element of this tale is false").

66. Look at the first six letters appearing at the top line of your keyboard.


68. *See Case C-418/01, IMS Health GmbH & Co. v. NDC Health GmbH & Co., 2004 E.C.R. I-5069, at ¶ 27–29* (addressing the argument made by IMS’s competitors that the use of the IMS structure has become indispensable in order to penetrate the market for pharmaceutical data, in light of users' dependence upon that structure). *See generally Part I.C supra* (discussing lock-in).

Undeniably, lock-in situations are not limited to copyright protected subject matter. After all, the ultimate lock-in example in economic literature—the QWERTY keyboard—is certainly not a copyright-protected subject matter. Yet, in other instances, the existence of copyright protection of selection and arrangement can influence the ability of others to create a convenient migration path to an alternative, competing database.

One interesting example is the legal-publishers’ battles concerning the arrangement of materials compiled and published under the West Reporter series. At the center of that litigation was the “star pagination” method employed by West’s competitors, under which the location of a page in the West series appears (marked by an asterisk) inside the competitors’ legal databases. This technique, probably familiar to most readers, allows users to cite cases in accordance with the West Reporter, even when they are using an alternative database. Thus, “star pagination” enables users to maintain the network advantages of the structure of the West database, while accessing the same information through a competing database. On several occasions, West contended the star pagination technique infringed its copyright in the arrangement of its compilations. Two different courts agreed with West in separate cases, first in 1986 and then in 1996. In both cases, the competing publishers, among them LEXIS, which has become a chief rival of West in the computerized legal database industry, were enjoined from employing “star pagination.” However, in 1998, publisher Matthew Bender succeeded in obtaining an opposite outcome. In that case, the Second Circuit Court of Appeals held, inter alia, that the pagination of the West cases is unprotected by copyright, and that “star pagination” does not infringe West’s copyright in its arrangement. The decision opened the door to the use of “star pagination” again.

70. The QWERTY keyboard was developed in the 19th Century, so any intellectual property rights that ever subsisted in it expired long ago. See generally supra notes 61–66 and accompanying text.
71. West Publ’g Co. v. Mead Data Cent., 799 F.2d 1219, 1227–28 (8th Cir. 1986).
72. As noted by one court, this practice is particularly useful in light of procedural requirements to cite the West version of certain federal decisions. See Matthew Bender & Co. v. West Publ’g Co., 158 F.3d 693, 696 (2d Cir. 1998).
73. Id.; West Publ’g Co., 799 F.2d at 1222; Oasis Publ’g Co. v. West Publ’g Co., 924 F. Supp. 918, 922 (D. Minn. 1996).
74. West Publ’g Co., 799 F.2d at 1227–28.
75. Oasis Publ’g Co., 924 F. Supp. at 924.
76. West Publ’g Co., 799 F.2d at 1229; Oasis Publ’g Co., 924 F. Supp. at 931.
77. Matthew Bender & Co., 158 F.3d at 699 (differing with previous decisions, the court held, inter alia, that West’s pagination did not entail even a “modicum of creativity” and was therefore an unprotected element of its compilation). But see id. at 710–11 (Sweet, J.,
pagination" by West's rivals. Assuming, for the purpose of the present discussion, that users prefer a database which allows references to the West Reporter, the example demonstrates that the scope of copyright protection of selection and arrangement can indeed affect users' choice and ability to switch between databases. More generally, in an environment of strong network effects, broad copyright protection of selection and arrangement can hinder users' ability to select a preferable database—a result which may harm social welfare.

In addition to its impact on economic lock-in, broad copyright protection of database selections and arrangements may increase the "mental lock-in" described in the previous sections. The control which copyright provides allows database owners to reinforce existing categories, classifications, and selections, and further deepens our perception of those structures as "given," rather than as relative and subject to challenge. Challenging prevailing beliefs and perceptions, on the other hand, can sometimes require certain use of those selections and arrangements. Consider again the arrangement of Google search results. Imagine an alternative search engine that ranks websites in accordance with gender equality principles. In order to make a point about gender perceptions, our imaginary gender equality search engine wishes to compare its ranking to Google's popularity-based ranking. This comparison would require repeated copying of the arrangement of Google search results. Such a meaning-making—or meaning-challenging—act might conflict with Google's copyright in its selection and arrangement of search results.

This hypothetical example demonstrates that overly broad copyright protection of selection and arrangement may deepen and strengthen existing power structures, and produce dialogic externalities. In this context, too, a caveat is in order: copyright law is certainly not the sole source of this problem. Social, cognitive, and communication analyses all indicate that categorization is an inevitable complexity reduction mechanism, which does not depend

---

78. See supra notes 49-50 and accompanying text.
upon the existence of copyright. That literature also clarifies that
the need to categorize sometimes results in difficulties to perceive
unique individual qualities, and forms bases to stereotyping, un-
founded generalizations, and prejudice. Limiting the scope of
copyright in selections and arrangements will certainly not provide
a magic solution to all such matters. Yet, in certain cases the con-
trol over selections and arrangements afforded by copyright law
can create an additional obstacle to challenging prevailing pat-
terns, hierarchies, and meanings. The limitation of that control, on
the other hand, can assist in overcoming those difficulties and ex-
pand our "universe of thinkable thoughts."

B. Selections and Arrangements v. Other Patterns and Formats

In light of this analysis, one should take a closer look at the pro-
tection granted by copyright law to database selections and
arrangements. Compilations are explicitly listed among the subject
matter of copyright under the U.S. Copyright Act. The Act explic-
itly provides that "compilations" include collections "of preexisting
materials or of data," which indicates that databases comprised of
facts or other non-protected materials can, in principle, attract
copyright protection. However, as noted earlier, the scope of the
protection granted to collections of facts has been fiercely debated,
and stood at the center of the Supreme Court decision in Feist,
which held that copyright protection does not extend to the factu-
ral data assembled in a compilation, regardless of the investment—
or "sweat of the brow"—expended in collecting the information. Rather, originality requires a minimal degree of "creativity," which
should be expressed in the compilation's "selection, coordination,
or arrangement." Applying the creativity standard, the court con-
cluded that the alphabetically arranged white pages telephone

80. See supra notes 18–25 and accompanying text.
81. See Allport, supra note 14, at 20–21; McGarty, supra note 13, at 24–27; cf Bovk-
er & Star, supra note 5, at 82 (noting that reality is more complex than any classification
can recognize).
82. See Dabney, supra note 53.
84. See id. § 101 ("A 'compilation' is a work formed by the collection and assembling of
preexisting materials or of data that are selected, coordinated, or arranged in such a way
that the resulting work as a whole constitutes an original work of authorship. The term
'compilation' includes collective works.").
86. Id. at 363. The phrase "selection, coordination or arrangement" follows the defini-
directory at the center of that litigation lacked the requisite originality and thus could be freely copied by competitors. 87

*Feist* was thus perceived by its supporters as freeing facts and information from the grasp of copyright law, and the scholarly debate which followed focused primarily on the protection—or rather lack of protection—afforded to databases' factual contents. 88 On the other hand, the protection of databases' selections and arrangements on the basis of "a minimal degree" of creativity, 89 attracted considerably less attention and controversy. Thus, a database's selection and arrangement possessing a "minimal degree of creativity," 90 is unanimously protected, both in the United States and elsewhere. 91 Indeed, copyright protection of compilation selection and arrangement on the basis of creativity (or "intellectual creation") has emerged as an international standard, set out in the principal international agreements and conventions in the field of copyright. 92

Since *Feist*, the "minimal degree of creativity" test has been applied in numerous subsequent cases concerning database selections and arrangements. Many of those easily crossed the "minimal degree of creativity" threshold set out by the Court in *Feist*. 93 Thus, for example, the selections and arrangements of various classified directories were held sufficiently creative for copyright purposes. 94

88. See, e.g., sources cited supra note 2.
89. *Feist*, 499 U.S. at 348.
90. Id.
91. A similar standard for protection of selection and arrangement was adopted by the EU *Databases Directive*, which provides that databases shall be protected by copyright if "the selection or arrangement of their contents . . . constitute the author's own intellectual creation." Council Directive 96/9, art. 3.1, 1996 O.J. (L 77) 20 (EC). Comparable criteria appear in copyright laws of additional jurisdictions. See, e.g., Canadian Copyright Act, R.S.C. 2010, c. C-42 ("[C]ompilation" means (a) a work resulting from the selection or arrangement of literary, dramatic, musical or artistic works or of parts thereof, or (b) a work resulting from the selection or arrangement of data . . .").
92. See The Berne Convention for the Protection of Literary and Artistic Works art. 2(5), Sept. 9, 1886, as amended on Sept. 28, 1979 ("Collections of literary or artistic works . . . which, by reason of the 'selection and arrangement' of their contents, constitute intellectual creations shall be protected as such . . ." (emphasis added)); WIPO Copyright Treaty art. 5, Dec. 20, 1996, 36 I.L.M 65 ("Compilations of data or other material, in any form, which by reason of the selection or arrangement of their contents constitute intellectual creations, are protected as such. This protection does not extend to the data or the material itself and is without prejudice to any copyright subsisting in the data or material contained in the compilation." (emphasis added)). Almost identical language appears in Article 10(2) of the GATT-TRIPS Agreement 1994. Agreement on Trade-Related Aspects of Intellectual Property Rights art. 10(2), Apr. 15, 1994, 33 I.L.M. 1197.
94. See, e.g., CCC Info. Servs. v. Maclean Hunter Mkt. Reports, Inc., 44 F.3d 61, 67 (2d Cir. 1994) (holding that a "selection and arrangement" of guide books containing projected
This broad protection granted to selections and arrangements—almost as a matter of course—is puzzling in light of the significant social costs highlighted earlier. It is even more perplexing in light of the law’s reserved approach toward other patterns and formats within the ambit of copyright subject matter, from software “structure-sequence-and-organization” (SSO), through forms and codes, to television formats. Part of the analysis in this Article is applicable with respect to such structures as well: they, too, can be filled-in with different contents and offer their users (or audiences) a set of conventions that define the relations between the form and its specific contents. They, too, can constitute langue and can form a basis to different instances of parole. However, when such other patterns and formats are concerned, copyright law is not oblivious to the difficulties entailed in broad protection, and adopts a cautious attitude: take, for example, the “structure-sequence-and-organization” of computer software, a term which resembles to a large extent the term “selection and arrangement” used in relation to databases. Like selections and arrangements, SSO refers to the structural elements of computer programs and the relations between these elements, beyond the level of the computer code itself. Yet, unlike selections and arrangements, the broad copyright protection that was initially granted to programs’ structure-sequence-and-organization in the famous case of Whelan Assocs. v. Jaslow Dental Lab, Inc. encountered intense criticism for values for used cars was sufficiently original); Key Publ’ns, Inc. v. Chinatown Today Publ’g Enters., Inc., 945 F.2d 509 (2d Cir. 1991) (finding that the selection of a Chinese-American restaurant’s directory merited copyright protection, although the claim was eventually denied due to plaintiffs failure to prove copying); William A. Graham Co. v. Haughey, 430 F. Supp. 2d 458, 466 (E.D. Penn. 2006) (holding that a simple “selection and arrangement” of an insurance agent database was protected); Am. Massage Therapy Ass’n v. Maxwell Peterson Assocs., Inc., 299 F. Supp. 2d 941 (N.D. Ill. 2002) (finding that the “selection and arrangement” of a massage directory was sufficiently original, although the claim was eventually denied, as the copying referred solely to the contents of the directory); see also supra note 37 (discussing the legal publishers’ cases); cf Open Source Yoga Unity v. Choudhury, No. C 03-3182 PJH, 2005 WL 756558 (N.D. Cal. Apr. 1, 2005) (finding that the “selection and arrangement” of 26 yoga positions known as “Bikram Yoga” can, in principle, attract copyright protection). But cf. Bellsouth Adver. & Publ’g Corp. v. Donneley Info. Publ’g, Inc., 999 F.2d 1436 (11th Cir. 1993) (denying protection for a “golden pages” directory).

95. See Whelan Assocs., Inc. v. Jaslow Dental Lab., Inc., 797 F.2d 1222, 1248 (3d Cir. 1986) (holding that copyright protection of computer programs extends to the programs’ “structure, sequence, and organization”).

96. See generally Shur-Ofry, supra note 69, at 536–37 (discussing the interactive attributes shared by popular patterns and formats of various types).

97. Whelan Assocs., 797 F.2d at 1239. The defendant’s software in that case (an application program) was held to infringe the plaintiff’s software, due to the similarity between the structure, sequence and organization (SSO) of both programs, and despite the fact that the programs were written in different software languages and their codes were not similar. Id. at 1242–48.
being over-expansive and burdensome, and subsequent courts adopted more restrictive protection criteria. Copyright law has also been reluctant to protect a user interface of a widespread software, which was held to constitute an unprotected "method of operation" in the renowned Lotus affair. Similarly, copyright law is unenthusiastic to protect cultural genres and television formats. The latter are sometimes described as warranting "thin copyright," protecting only against literal or very close copying.

While the dynamic attributes of such other patterns and structures are not always explicitly acknowledged in case law, both scholars and courts seem to recognize that broad copyright protection of such subject matter entails difficulties, and the appropriate scope of such protection is often questioned or debated. Yet, when database selections and arrangements are concerned, such discourse is almost completely absent.

This inconsistency is even more apparent when considering copyright's cautious treatment of codes used in various industries and "blank forms." The latter are sometimes declared "systems"
or "methods of operation," which are denied any copyright protection. 105 However, the distinction between "systems" or "methods of operation," which are excluded from copyright protection altogether, and compilations' selections and arrangements, which are often afforded such protection, is far from clear. Both types of subject matter concern selection, editing, arrangement, classification, and categorization of information. 106

One can only assume that where "naked formats"—namely, structures that are not filled with particular data or contents—are concerned, the network value and dynamic attributes of the structure and the difficulties in granting it full copyright protection are easier to recognize. But when the subject matter is a database filled with valuable data and factual information—in itself unprotected by copyright—denying copyright from its selection and arrangement as well is perceived as too severe an outcome. In such cases, the tendency is to protect the database's format by describing it as selection and arrangement. It is also possible, that "copyright's obsession with truth," to borrow a phrase coined by Jane Ginsburg, 107 overshadowed the difficulties associated with copyright in selection and arrangement. In other words, the fierce debate over copyright protection of the facts and information compiled in databases created the impression that the protection of selection and arrangement, as opposed to the former aspect, does not entail any significant social costs. Both of these assumptions are reflected in the following reasoning of the Court of Appeals for the Second Circuit:

Without financial incentives, creators of such useful compilations might direct their energies elsewhere, depriving the public of their creations and impeding the advancement of learning.

The grant of such monopoly protection to the original elements of a compilation, furthermore, imposes little cost or

105. 17 U.S.C. § 102(b) (2006). This principle is often attributed to the decision of Baker v. Selden. See Samuelson, supra note 6, at 207.

106. See Southco, 390 F.3d at 289–90 (Becker, J., concurring) (wondering whether the plaintiff's part numbers are "a compilation of data, a system of classification, or something else"); cf. Karjala, supra note 6, at 195–97 (highlighting the difficulties to distinguish between compilations, valuations, taxonomies, codes, systems, and methods).

disadvantage to society. The facts set forth in the compilation are not protected and may be freely copied; the protection extends only to those aspects of the compilation that embody the original creation of the compiler.108

It is difficult to dispute the court's observation regarding the financial incentive required in the database industry. Despite the difficulties highlighted throughout this Article, the social contribution of well-organized access to information is undoubted, particularly in an age of information overload.109 Creators of original selections and arrangements certainly need—and deserve—the economic incentives provided by copyright, particularly under the post-*Feist* regime, which leaves databases' factual contents unprotected. Indeed, recent scholarship has cast some doubt on the significance of monetary incentive in promoting innovation.110 However, the traditional incentive narrative does seem to prevail with respect to databases and compilations, due to their utilitarian-commercial nature.111

I am not arguing, therefore, that database selections and arrangements should be deprived from copyright protection altogether.112 Nor am I arguing that such selections and arrangements should be automatically declared unprotected “systems” or “methods.” Instead, I call for solutions that strike a more nuanced balance between the need for incentive and the social costs of copyright protection. The aim of this Article, however, is to highlight the other side of the equation, namely the social costs entailed in broad copyright protection of selections and arrangements, which must be taken into account when searching for that ever elusive balance.


109. *See* discussion *supra* Part II.A; *see also* Berring, *supra* note 50, at 316–18 (highlighting the significance of organizational thinking in the information age).


111. *See* Yen, *supra* note 2, at 1374–76 (analyzing the significance of incentives for the production of databases).

112. *Cf.* Karjala, *supra* note 6, at 174 (arguing that protection of compilations' 'selections and arrangements' under the creativity threshold is required in light of the wording of Section 102 of the Copyright Act).
By so doing, the analysis in this Article challenges the additional observation made by the Maclean court:113 it indicates that copyright’s disregard of the attributes of selection and arrangement does have a price. Overly broad protection of this element may enhance the externalities associated with categorization.114 Furthermore, it damages the coherence of copyright law, by the unequal treatment of this paratext, in comparison to other formats included in the ambit of copyright protected subject matter. It may also be that the actual difficulties entailed in protecting certain selections and arrangements lead some courts to deny such protection, using “lack of creative spark” as a justification (or rather an excuse), while increasing ambiguity and incoherence as to the level of creativity actually required in copyright law.115

C. Looking Ahead

The above analysis demonstrates that the topic of database selection and arrangement should be marked as an issue that requires a more cautious approach from copyright’s perspective. One possible solution would be to calibrate the fair use doctrine to handle problematic cases concerning selections and arrangements. Thus, for example, the doctrine could apply in circumstances where a protected selection and arrangement is used for criticizing existing social structures and perceptions.116

Another potential direction may be to focus particularly on those selections and arrangements that become industry standards. The discussion in this Article demonstrates that the difficulties

113. See supra note 108 and accompanying text.
114. See discussion supra Part II.A.
115. Thus, for example, it is difficult to avoid the impression that the reluctance to protect the West pagination was influenced by the dominance of West’s ‘selection and arrangement, even if formally justified by the absence of a minimal degree of creativity. See Matthew Bender & Co. v. West Publ’g Co., 158 F.3d 693, 696 (2d Cir. 1998) (holding that the pagination of the West case reporter lacks sufficient creativity); Matthew Bender & Co. v. West Publ’g Co., 158 F.3d 674, 682 (2d Cir. 1998) (holding that West’s ‘selection and arrangement’ of judicial opinions lacks a creative spark, since it is dictated by an industry standard). But cf. id. at 692–93 (Sweet, J., dissenting) (noting that the industry standard was created due to the success of the plaintiff’s work); Kajala, supra note 6, at 192–200 (arguing that the creativity test has a blurring effect and leads to overly broad protection of various works that are not compilations).
116. See supra notes 78–79 and accompanying text. The Fair Use doctrine comes to the assistance of numerous defendants who attempt to challenge social conventions by using copyright protected works. Two famous examples are Suntrust Bank v. Houghton Mifflin Co., 268 F.3d 1257 (11th Cir. 2001) (a challenging adaptation to Gone With the Wind) and Mattel, Inc. v. Walking Mountain Prods., 353 F.3d 792 (9th Cir. 2003) (challenging the depiction of ‘Barbie’ in an art installation).
entailed in broad copyright protection of selections and arrangements are especially prominent when the databases in question acquire a standard status. These databases are more likely to yield cognitive dependency as well as economic and mental lock-ins. Several commentators have recently proposed an expiry of copyright in utilitarian standards. Elsewhere I have argued that the doctrine of copyright misuse can become a useful tool in handling the difficulties associated with copyright protection of standards across the entire range of copyright protected works. This doctrine may be particularly suitable with respect to the selections and arrangements of standard databases, as it enables a range of flexible and more nuanced outcomes and does not necessitate the complete expiry of copyright.

And lastly, one may also consider rethinking the current balance in copyright law between the lack of protection of database contents and the broad protection of database selection and arrangement. If, indeed, as speculated earlier, overly broad protection of database selections and arrangements sometime compensates—in a very inexplicit and inadvertent manner—for the complete lack of protection afforded to factual contents, then one may consider changing the equation by affording certain protection to assembled factual contents, while narrowing the scope of protection afforded to selections and arrangements. Yet, that speculation certainly requires further support and necessitates careful consideration of the issue of copyright protection of factual contents, an issue which is beyond the scope (or perhaps outside the selection?) of this Article.

In any case, while this Article does not purport to offer a complete doctrinal solution to the problems associated with databases' dynamic dimension, highlighting these aspects can hopefully form a starting point for a more detailed discussion of the appropriate solutions.

117. Indeed, many 'selections and arrangements' cases concerned such "standards." See, e.g., Matthew Bender & Co., 158 F.3d at 692-93 (Sweet, J., dissenting) (discussing West's 'selection and arrangement' which has become an industry standard); CCC Info. Servs., Inc. v. Maclean Hunter Mkt. Reports, Inc., 44 F.3d 61 (2d Cir. 1994) (discussing popular guide books for used vehicles' projected values, reference to which was mandatory under insurance regulation); Open Source Yoga Unity v. Chodhury, No. C 03-3182 PJH, 2005 WL 756558, at *9 (N.D. Cal. Apr. 1, 2005) (discussing the subsistence of copyright in the standard arrangement of "asanas" under the "Birkam Yoga" school ); see also supra Part IA (discussing the European IMS case, in which the subsistence of copyright in the standard arrangement led to antitrust proceedings in the European Court of Justice).

118. See, e.g., Koelman, supra note 6; Samuelson, supra note 6.

119. Shur-Ofry, supra note 69, at 574–77.
CONCLUSION

This Article explored the dynamic narrative of the copyright-databases relationship. While the prevailing account of that relationship focuses on access to information, this Article demonstrated that databases are not stagnant piles of information. Rather, they play an important social role in defining relations, hierarchies, and interactions between pieces of information. By including, excluding, structuring, contextualizing, indexing, classifying, and categorizing information, databases fulfill a function which is necessary for human cognition and interaction.

This dynamic dimension highlights cognitive and economic dependence upon familiar structures and clarifies the inclination of databases towards standardization. These attributes can yield economic and mental lock-in, which overly broad copyright protection might deepen and increase. On a normative level, databases—even the most technical, mundane ones—are sites fraught with inevitable judgments, tensions, and power relations, that define and shape our "universe of thinkable thoughts" in manners difficult to perceive and to challenge. From an economic perspective, if databases are susceptible to user lock-in, that may harm social welfare. Notably, these externalities do not stem from particular facts or data stored in databases; rather, they concern the structures and scope of such compilations.

The analysis in this Article shifts the focus of the copyright-databases relations from the protection of contents to the protection of selection and arrangement, an aspect which has been almost taken for granted by both courts and scholars. It illustrates that broad control over selections and arrangements afforded by copyright law can increase the difficulties and externalities associated with the structural-dynamic attributes of databases. This analysis is not limited to physical, off-line databases. It also applies, to a large extent, to computerized online databases that govern our networked world. Indeed, the social costs entailed by the protection of selections and arrangements may be greater than those entailed by the protection of databases’ factual contents. Clearly, they should not be clouded by the ongoing debate over the latter aspect. Overlooking the issue of copyright in selection and arrangement not only burdens the handling of those difficulties, but it also harms the integrity and coherence of copyright law.

120. See discussion supra Part I.B. For the image of "universe of thinkable thoughts," see Dabney, supra note 53 and accompanying text.
121. See discussion supra Part I.C.
By highlighting the dynamic dimension of databases, this Article calls for a more cautious and conscious approach toward copyright protection of selections and arrangements. It further hopes to form a starting point for further discussion that will shift at least part of the focus of the copyright-databases debate from access to information to selection and arrangement.