Peer to Peer Meets the World of Legal Information: Encountering a New Paradigm

Ethan Katsh  
ethan.katsh@nyls.edu  

Beth Simone Noveck  
New York Law School

Follow this and additional works at: http://digitalcommons.nyls.edu/fac_articles_chapters  
Part of the Intellectual Property Law Commons

Recommended Citation  
Peer to Peer Meets the World of Legal Information: 
Encountering a New Paradigm*

Ethan Katsh** and Beth Noveck***

The authors describe a proposed system for patent application reviews that uses new technologies to access information—community peer reviews. By allowing examiners to “mine for data” in the heads of experts rather than in libraries or databases, the proposal illustrates how new technology could change the boundaries of legally authoritative and relevant information and make it possible to identify legitimate authority from new sources.

§1 The first significant introduction of information technology into the world of law, more than two decades ago, was through the library. While automation technologies had been used to organize collections prior to that, it was then that machines first began to change the experience of obtaining legal knowledge by making it possible to access information located far away. Since then access to information has become faster, easier, and cheaper, each by an order of magnitude. But new technologies do not simply provide more powerful, efficient, and convenient methods for extracting knowledge from traditional sources of legal authority. Instead, the new communications technology that connects people across networks may make it possible to identify legal authority from new sources that were not previously relied upon or perceived as legitimate.

§2 Robert Berring has written that “[l]egal information is in the midst of great change, a change not just in formats, but in the authority structure of the materials that legal workers use. A redefinition of the most basic sort is taking place.”1 By making authoritative information not only widely available to consume but widely available to produce, technology is potentially enabling groups of the human knowledge of ordinary people’s experience, rather than just books produced by an elite few, to become sources of authority. We have always had institutions like the jury that looked to people outside the profession as a source of legal authority, and the Administrative Procedure Act has mandated citizen participation in agency rule

* © Ethan Katsh and Beth Noveck, 2007. This article is an edited version of remarks delivered at Legal Information and the Development of American Law: Further Thinking about the Thoughts of Bob Berring, a symposium held at Boalt Hall on the University of California, Berkeley campus, Oct. 21, 2006.

** Professor of Legal Studies and Director of the Center for Information Technology and Dispute Resolution, University of Massachusetts at Amherst, Amherst, Massachusetts.

*** Professor and Director, Institute for Information Law & Policy, New York Law School, New York, New York; McClatchy Visiting Associate Professor, Communication Department, Stanford University, Palo Alto, California.

making for the last sixty years. Yet, for the most part, legal authority has been codified from the work of specific members of the legal profession—be they lawyers in private practice, bureaucrats in government, or judges in courts—operating within the bounds of a limited set of professional institutions.

§3 Now the citizen journalism movement challenges who has the authority to be called a journalist. Medical Web sites written by patients erode the perception that only licensed health care workers possess the requisite knowledge and expertise to heal. Similarly in law, the rise of social networking technology could fundamentally transform the assumption that legitimate authority comes only from books.

§4 The success of endeavors such as Wikipedia (www.wikipedia.org) or Google Answers (http://answers.google.com/answers) demonstrates that there are other sources of authority beyond those contained in published works. The array of informational riches that the Internet puts at our disposal, including not only a wider array of texts, but also lived, human experience, reveals the meager paucity of the informational diet on which our legal institutions currently subsist and upon which their legitimacy precariously relies.

§5 Identifying all the types of decisions and contexts in which people working together online might provide a legitimate source of counter-authority to traditional legal texts must evolve over time. But we can start with examining one example of the shift to new sources of legal authority for legal information by looking at a proposal made by one of the authors of this article and adopted by the United States Patent and Trademark Office in the area of patent examination.\(^2\) This “Peer to Patent” idea reflects a radical shift in the location and valuation of legally relevant information. It will help to demonstrate that it is possible to rely upon dynamic, human expertise from a wider array of individuals rather than upon codified and accreted knowledge from licensed professionals even in contexts, unlike Wikipedia, where decisions are important and have precedential value.

**The Patent Process and the Informational Challenge**

§6 A recent report from the U.S. Copyright Office noted that “[u]ntil the late 1990s copyright was more or less invisible to the general public.”\(^3\) In the last few years, however, copyright issues have become of central concern and importance to a newly media-savvy public. Today, patents are “the new black.” Until the last few years, the Patent Office and the manner in which patent applications were processed were even less in the public eye than copyright issues had been. This has changed, however, as the number of patent applications has increased and as questions have been raised as to how and why the overwhelming majority of patent applications were approved.

---


A successful patent application must meet five major statutory criteria: patentable subject matter, utility, novelty, nonobviousness, and enablement. The core of the examination centers around two questions: is the invention new and is it enough of an advance over what came before as to be not obvious to those with expertise in that area of invention. To answer both questions, the patent examiner must search for “prior art,” know-how that predates the invention and might shed light on whether the claimed invention differs from previous inventions and whether the invention was in public use (e.g., on sale or published) by anyone else, including the inventor.

In their recent book, Innovation and Its Discontents, Adam Jaffe and Josh Lerner document many patents that are anything but novel and nonobvious, such as patent number 6,368,227 for “Method of Swinging on a Swing” awarded to a five-year-old boy (subsequently cancelled). Or patent number 6,574,645, a patent on a method for drafting a patent. The patent awarded to Smucker’s for the crustless peanut butter and jelly sandwich is, by now, legendary. Jaffe and Lerner describe a patent sought for expirationless options thirty years after economists won the Nobel Prize for the same idea.

A key problem in reviewing patent applications is that, as old and obvious as an idea might seem, finding the relevant prior art to invalidate it is often quite difficult, especially considering the time and resource constraints under which patent examiners work. This creates an opportunity to exploit deficiencies in the system and to undercut belief in the legitimacy of the process. The United States Patent Office now receives nearly four hundred thousand patent applications per year and has a backlog approaching a million. Between 1990 and 2004, the number of patents issued in the United States nearly doubled.

---

6. U.S. Patent No. 6,368,227 (issued Apr. 9, 2002) (“Method of Swinging on a Swing”), noted in JAFFE & LERNER, supra note 5, at 34.
8. See JAFFE & LERNER, supra note 5, at 25–26, 32–34.
¶10 When examining a patent application, under current law, patent examiners may consult databases to search for prior art, but they may not consult individuals or request information from the public. Examiners are expected to be scientifically adept enough to discover the prior art on their own. After eighteen months, most applications are published; once published, it is possible for the public to submit written evidence of prior art. Though there is a backlog of one million applications, last year the Patent Office received only between forty and a hundred written public comments. The examiner is still forbidden to engage in external communication. The public is even prohibited from providing commentary or explanation with the prior art submitted. Furthermore, the patent applicant is not required to submit prior art of his or her own along with the application. The entire informational burden rests with the patent examiner.

¶11 Trying to determine whether or not a process or invention is novel poses an interesting informational challenge in that it is not really possible to search for novelty. A conclusion that novelty exists is a determination that depends not on finding something but on not finding something. For someone seeking a patent, a “failed” search is what is desired since a failed search shows that the idea has not been practiced before.

¶12 One conducting legal research is most often engaged in an effort to find missing data that will either support or not support some proposition. Generally, whether a particular search is successful or not can be judged by whether the information one is looking for has been found. When one conducts a search in the hope of not finding something, such as a literature review conducted to determine whether to undertake some research project, the value of the search depends less on what one finds than on the quality of the search process, on whether the tools
used and the strategy followed allow one to conclude that one has found every-
thing that could be found. In the Patent Office context, therefore, for a “novelty”
assessment to have legitimacy, it must not only be a failed search but a compre-
hensive search.

¶13 In practice, the examiner searches an internal Patent Office database con-
taining primarily U.S. and foreign patent applications in an effort to find prior art. The examiner might also consult databases containing a limited number of jour-
nals. In some cases, when office policy permits (and it doesn’t do so in all divisions of the Patent Office where there is a concern about security breaches resulting from Internet searching), an examiner might use Google or Slashdot but is limited to those sources that can be found on his or her own from the office. In searching for prior art to invalidate a patent’s claims, the examiner sometimes turns up nothing. While the patent may sound like something familiar that has come before, often the examiner cannot find other written material that actually “teaches the claims” of the patent directly. An invention might be in an area of innovation, such as computer software, where much of the information is not published in patents or journals. Alternatively, the patent examiner is inundated with related prior art but has trouble in the time allotted to review an application, winnow the material and find art that is relevant to determining patentability.

¶14 Even if pertinent prior art is found, the examiner still may have trouble knowing, from the perspective of one working in that area of science, whether the patent is an obvious or nonobvious inventive leap over the combined prior art references. The search for prior art is aimed not only at assessing novelty but also at determining whether the invention represents a significant enough advance over what came before. The “obviousness” determination inquires as to whether “the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.”13 As patent scholar Rebecca Eisenberg explains:

[T]his language seems to call for evaluations of nonobviousness from the perspective of ordinary practitioners who are contemporaries of the inventor in the relevant technological community. It specifies a point in time as of which the obviousness of the invention should be evaluated (“at the time the invention was made”) and designates the person whose judgment of obviousness should control (“to a person having ordinary skill in the art to which said subject matter pertains” or PHOSITA), as well as directing attention to “the differences between the subject matter sought to be patented and the prior art.”14

¶15 The examiner’s manual suggests three bases for assessing obviousness. There must be “some suggestion or motivation” available to one of ordinary skill

in the art, "to modify the reference or to combine reference teachings."15 In other words, there needs to be evidence that one with some expertise in this arena would know and want to make the modifications that would produce the jump from what came before to what is claimed in the application. Second, "there must be a reason­able expectation of success" in making that leap.16 Finally, the prior art, whether a single item of art or the existing literature when combined, "must teach or suggest all the claim limitations."17

16 It is essential to this determination that the examiner put him or herself in the shoes of the fictional person skilled in the art and render the decision through that person’s eyes and not his or her own. This requires identifying this person and characterizing his skill set.18 The person skilled in the art is considered to be the best metric for the scope, content, and meaning of prior references. But the ultimate determination of patentability is a legal one.19 While the person skilled in the art could bring training, judgment, intuition, practical knowledge, and skills in the craft, the examiner may not consult such persons directly and, instead, must, at present, be content to substitute his or her judgment,20 informed by the limited set of electronic information resources to which he or she has access.

17 Under the current patent process, the “expert” with the requisite knowledge to make the decision is the patent examiner. Our intellectual property law and administrative practice have been constructed around the belief that centralized administrators have the best access to information, that expert bureaucrats are the only way to produce dispassionate decisions, and that making decisions in the public interest requires keeping the public at bay. At one point in time, the Patent Office was a premiere repository of scientific information. Given that public consultation has been difficult and time-consuming, it made sense to build legal institutions around this insular bureaucratic conception of expertise. But that is premised upon a bygone material reality.

A New Approach to Legal Information:
The Community Patent Review Project

18 The informational problem faced by the Patent Office is one that can only grow with the increasingly complex and challenging nature of innovation and,

16. Id.
17. Id.
18. See Graham v. John Deere Co., 383 U.S. 1, 17-18 (1966) ("Under § 103, the scope and content of the prior art are to be determined; differences between the prior art and the claims at issue are to be ascertained; and the level of ordinary skill in the pertinent art resolved. Against this background, the obviousness or nonobviousness of the subject matter is determined. Such secondary considerations as commercial success, long felt but unsolved needs, failure of others, etc., might be utilized to give light to the circumstances surrounding the origin of the subject matter sought to be patented. As indicia of obviousness or nonobviousness, these inquiries may have relevancy.").
19. Id. at 17 ("[T]he ultimate question of patent validity is one of law... ").
20. See Eisenberg, supra note 14, at 888.
therefore, of patent applications. It is not simply that increases in the number of patent applications filed will require more searching by more examiners. More significant is that potentially relevant information is located in a broadening array of locations. We are in an age in which knowledge production and acquisition are accelerating while the kinds of online sites in which knowledge is being stored are proliferating. As a result, searches of those sources of information that traditionally revealed whether prior art existed may now reveal a larger quantity of information but may also appear less and less comprehensive. Tools for searching for information online have improved, but that does not necessarily help if the key concern is being able to conclude confidently that everything that could be found has been found.

¶19 The Community Patent Review Project is an initiative of the New York Law School Institute for Information Law & Policy in collaboration with the United States Patent and Trademark Office that aims at improving the quality of issued patents by providing patent examiners access to better information through an open network for community peer review of patent applications. In a manner of speaking, the proposal aims to reform the patent examination system by allowing patent examiners to "mine for data" in the heads of experts rather than in libraries and databases. Examiners would still not consult the public directly (as this would require statutory reform) but would allow the public to submit prior art to the Patent Office. Yet, this public consultation software system for patents would allow information to come in from a variety of sources wider than the limited databases to which examiners now have access. The proposal builds upon the notion that expertise is not centralized but distributed in the minds of those with the requisite knowledge. Such a proposal can be implemented because of the development of software applications intended for collaborating, sharing, and aggregating data. The goal of such a system would be to help the patent examiner find the right references and have access to those who can advise on how to combine them, akin to having expert witnesses before-the-fact during examination. This might help to introduce better information into the process and thereby winnow bad patents, narrow the claims of good patents, and improve the quality of those inventions awarded a patent.

¶20 Technology is providing the opportunity to move away from a model of research done in a traditional way, by an individual searching sources of published information. An alternative, or an addition, to individualized searching is a vision of collaborative expertise whereby the know-how of a large, trained, and dedicated governmental staff with legal expertise can be harnessed to the wisdom of those with deep scientific, subject-matter expertise. Using communication technology, it is possible to create a new mechanism for large-scale distributed decision

---

making\textsuperscript{22} that distinguishes legal from scientific decisions. With procedures in place to distribute but interconnect these two forms of expertise, it is possible to create new mechanisms for making administrative decisions more broadly. The idea of scientific citizen juries, blue ribbon panels, or advisory committees\textsuperscript{23} is not new. But the suggestion to use newly available social reputation software—think Slashdot karma or eBay reputation points—to make such panels big enough, diverse enough, democratic enough, and trustworthy enough to assist the patent examiner is new.

\textsuperscript{21} In August 2006, the United States Patent and Trademark Office announced that it would pilot just such an open peer review system as part of its strategic initiatives for the forthcoming years.\textsuperscript{24} It is the first time that a lawmakers has attempted to leverage this kind of human expertise using social software. Of course, many government agencies consult with scientists in awarding grants, with experts in making policy, and even with the public when drafting rules. But the grant process is done using closed peer review panels for which the agency selects the experts; they do not select themselves. Similarly, when agencies seek to reach out to scientists, it is through closed invitation-only processes. Public consultation in rule making, while more open and even available via the Web, has had little real impact on decision making. Agencies frequently outsource the reading of comments to third parties. In no event do any of these practices leverage "social software" using technology to create a network of self-selected and mutually self-rating experts.

\textsuperscript{22} The Community Patent Review process, by contrast, offers an open Web site where, with the benefit of available educational materials, users could submit prior art and commentary in response to published patent applications. Participating reviewers could use the original software not only to submit bibliographic information, but also to rate and rank that information. In so doing, they would collaboratively create a rank-ordered list of citations. The software would then forward only the top ten citations to the patent examiner for review. This process would be both accessible and practical to use for reviewers and examiners. The software also allows participating reviewers to rate each other on the basis of their expertise and usefulness in the process. Expertise would flow from how the reviewers perform as information providers to the Patent Office, not from the degrees or qualifications they have earned outside the system.

\textsuperscript{23} Federal Advisory Committee Act (FACA), 5 U.S.C. app. \textsection 2 (2000) (authorizing the establishment of a system governing the creation and operation of advisory committees in the executive branch).

¶23 This proposal for changing the process for locating prior art is aimed at narrowing the gaps in the patent system’s filter: it may increase the likelihood that good inventions will pass through while blocking unworthy inventions. The proposed system directly addresses the problems with the current examination process by enabling the community of practice to collaborate on finding prior art and transforming the “person skilled in the art” from a fictional legal personage into a real group. This system would augment review by a lone examiner with assistance from experts in the relevant area of art. While the system requires little to no statutory or regulatory change, it would precipitate an overhaul in the way patents are examined.

The Jurisdiction of Print and Text

¶24 The new source of information that technology is opening up for us is actually the oldest of possible sources: the expert information that lies in the minds of human beings. Individual human experts do surface occasionally in the legal process, such as in juries or in providing expert testimony as witnesses at trials. In most searches for legally relevant information, however, the opinion or perspective of individual humans is not sought. In recent centuries, authoritative legal information has been considered to be located not in unfiltered information provided by a person but in information that has been filtered and then placed in an accessible medium. In other words, it is not what experts are actually thinking that is sought and trusted, but what they have thought and what others have then considered and preserved.

¶25 Before the intrusion of digital networks, the world of accessible and relevant legal information was bounded by the physical walls of the library. This was understandable and resulted from the efforts of librarians to define and create “collections” filtered by editors. Even judicial opinions, content that clearly belongs in a legal collection, went through a correcting, if not an editing, process.25

¶26 Legal research systems rarely sought or engaged information outside of law library collections. That would have implied that a particular law collection was

25. It is, I (Ethan) think, worth noting here, even though it is not consistent with the impersonal writing style one employs in articles like this, that my first thought upon writing this sentence was that I needed to find some published source to use as a footnote. I realized quickly that there was some irony to this in that it illustrated how strong our training is to favor the published over something personally known or experienced. The basis for the comment about judicial opinions is knowledge that I have as a result of a visit some years ago to what was, at the time, the West Publishing Company. I probably could have put this fact in a footnote if I could not find a published source but if I were to simply indicate in a footnote that my source was myself, I could also easily imagine receiving a note at some point from a law review editor that I really should be looking for something published, that my memory was not as trustworthy as something published. Of course, what would be perfectly acceptable would be a citation to an article I might have written in which I had described my trip to West. I think I have made my point but if you do not trust my memory, see Deborah Tussey, Owning the Law: Intellectual Property Rights in Primary Law, 9 Fordham Intell. Prop. Media & Ent. L.J. 173, 181–83 (1998).
lacking, and, because most law libraries are physically separate from other libraries in a university or town,26 also would have required traversing some physical distance. Only rarely would limiting one’s search exclusively to the law library be deemed insufficient or incomplete because outside sources were not consulted.

¶27 As legal information that was previously in the library moved online, it could be searched more rapidly and more efficiently than before, but the physical guideposts that influenced our understanding of where legally relevant information could be found (and, equally important, where it could not be found) started to erode. Law collections can now be accessed by persons outside the law library and those in the law library have increasing access to materials that were previously not part of law collections. As a result, we have more and more information at our fingertips and more and more capabilities to access that information, but the task of conducting a thoroughly comprehensive and complete search, and of defining what a law “collection” is, has become more challenging. In other words, when all legally relevant information was in the library and the tool of choice was a West digest, supplemented perhaps by the Index to Legal Periodicals, a few other bibliographical tools, and a reference librarian, one was probably more certain than now that one had found everything relevant that could be found.

¶28 More than a century ago, Christopher Columbus Langdell asserted that “law is a science, and . . . all the available material of that science are contained in printed books. . . . [L]aw can only be learned and taught in a university by means of printed books. . . . [P]rinted books are the ultimate sources of all legal knowledge.”27 This was a provocative and debatable claim but also one that suggests how powerfully print has touched both our minds and our practice. Legal research was not only to be located within law libraries, but print itself, as Berring has argued, provided the law with a supporting conceptual infrastructure on top of which could be built traditions and practices as well as physical structures.28

¶29 Maintaining content while placing it in a different medium may provide a variety of efficiencies but it also leads to a weakening of these supports. Thus, as legal research, quite visibly, acquired new tools, the universe of relevant legal information began, albeit less visibly, to grow larger and to have more ambiguous boundaries. In the age of print, as noted earlier, judicial opinions were enhanced and organized before being published. In the age of electronic access, controlling access to information became more difficult and users learned that there were “unpublished opinions,” a body of information that was easily excluded in the print environment and a practice that was sustained through a filtering process that most lawyers were unaware of.29

27. Christopher Columbus Langdell, Address at Harvard University “Quarter-Millennial” Celebration (Nov. 5, 1886), in 3 L.Q. REV. 123, 124 (1887).
Peer to Peer Meets the World of Legal Information

§30 The culture of print, as well as the technology of print, has been weaving its way through law for the last five centuries. The technology of print, in the centuries after Gutenberg, provided new tools and a new information infrastructure on which new legal information practices could be based. The culture of print ascribed value to activities employing these tools and resources. An emerging system of precedent, for example, was made possible by new tools for storing and accessing information, tools, unlike the written word, that could be trusted to accurately contain the words of the author. Printing's most frequently noted benefit was an increase in the number and copies of books, but the authority of the printed word is due to other qualities, namely, the uniformity of all printed copies and the implicit acceptance of the fact that what appears on paper is what the author actually wrote. Today, information in print receives an aura of trust and authority simply by being in printed form. The degree of authority can vary depending on the reputation of the author, the persuasiveness of the argument made, and the publisher, but this value comes on top of the value derived from the fact that the work has been printed. While we are frequently reminded not to "judge a book by its cover," the authority of print is due in some part to the fact that we do assign value to books with certain kinds of covers and certain images of what is between the covers.

§31 It has been written that the Internet is "perhaps the biggest collective expert ever known to humanity." We are in the process of acquiring tools to tap into this intelligence in ways that were not possible before and, as a result, not thought about before. What is being proposed is not the same as consulting a bigger library or surveying a panel of experts. A survey is a familiar method for obtaining information that information technologies have made more efficient. What is proposed here is not simply to communicate more efficiently with a group but to rely on information processing in a new way.

§32 The Community Patent Review proposal is important for what it suggests about the future as well as what it might contribute to the present. It is, most obviously, one of many examples of new technologies that allow access to information that was not available before. More importantly, perhaps, in the long run, it is an example of encounters with novel resources that may be changing the boundaries of legally authoritative and relevant information. When they were first revealed, the unpublished opinions of judges represented a new information source that placed the law in an awkward and uncertain position. These opinions came from an authoritative source, but this source did not want the opinions to have authority. Over time, some resolution of the quandary has occurred. The collective wisdom that is at the core of this "peer to peer" patent proposal can also be considered an


unpublished opinion. Indeed, that is exactly what it is. It is an opinion emerging not from an individual but from a group assisted by newly available information-processing capabilities.

Print allowed sources that had originated far away to retain the authority that previously required physical contact or acquaintance with the author. The new technologies have acquainted us with new communication mechanisms and have prepared us for the delivery of expertise as well as authority from afar. The peer-to-patent proposal moves us beyond using the network for communicating or transporting messages. The expertise relating to "prior art" that could be presented to the patent examiner relies on machine-based information processing as well as on human expertise and thus is a model for delivering expertise from afar, not only more efficiently than in the past but in a manner that is quite different from how it was delivered in the past. We are, as a result, presented with an interesting challenge, namely, to think creatively about how to match new capabilities for generating information to particular legal processes that are in need of improvement. Librarians, in particular, are likely to encounter new sources of knowledge that will be available, not because they exist somewhere and are brought to us, but because they are created in ways that were not possible with the tools we have been employing. Technology has placed us in a situation where, as Professor Berring, in whose honor this article has been written, has stated, "[w]e need a new set of thinkable thoughts."32

32. Robert C. Berring, Legal Research and the World of Thinkable Thoughts, 2 J. APP. PRAC. & PROCESS 305, 314 (2000). Professor Berring attributes the phrase "thinkable thoughts" to Dan Dabney. Id. at 311 n.13.