

WHAT CAN INFORMATION TECHNOLOGY DO FOR LAW?

Johnathan Jenkins\*

TABLE OF CONTENTS

I. INTRODUCTION.....	589
II. INCENTIVES FOR BETTER INTEGRATION OF INFORMATION TECHNOLOGY AND LAW.....	591
III. THE CURRENT STATE OF INFORMATION TECHNOLOGY IN LEGAL PRACTICE.....	594
IV. THE DIRECTION OF LEGAL INFORMATICS: CURRENT RESEARCH.....	597
<i>A. Advances in Argumentation Models and Outcome     Prediction.....</i>	<i>597</i>
<i>B. Machine Learning and Knowledge Discovery from     Databases.....</i>	<i>600</i>
<i>C. Accessible, Structured Knowledge.....</i>	<i>602</i>
V. INFORMATION TECHNOLOGY AND THE LEGAL PROFESSION: BARRIERS TO PROGRESS.....	604
VI. CONCLUSION.....	607

I. INTRODUCTION

MUCH CURRENT LEGAL WORK IS EMBARRASSINGLY, ABSURDLY,  
WASTEFUL. AI-RELATED TECHNOLOGY OFFERS GREAT PROMISE TO  
IMPROVE THAT SITUATION.<sup>1</sup>

Many professionals now rely on information technology (“IT”) to simplify, automate, or better understand aspects of their work. Such software comes in varying degrees of sophistication: less sophisticated tools include word processors, e-mail and instant messaging systems, file servers, and the like, while more sophisticated tools reach into the analytical core of a professional’s work. Although modern law firms and courts are awash in these less sophisticated tools,

---

\* Harvard Law School, Candidate for J.D., 2008. Special thanks to Paul O’Connell and Avi Pfeffer for fostering an interest in some of the topics discussed in this Note; and thanks to Doug Kochelek, Richard Heppner, and the *Harvard Journal of Law & Technology* Student Writing Committee for their many helpful comments on earlier drafts.

1. Marc Lauritsen, *Artificial Intelligence in the Real Legal Workplace*, in INFORMATION TECHNOLOGY AND LAWYERS 165, 175 (Arno R. Lodder & Anja Oskamp eds., 2006).

legal practice often lacks analogues to the more sophisticated tools found in many other industries.

A broad variety of industries have incorporated sophisticated data-manipulation techniques in recent decades. High-profile examples include the use of statistical data-mining techniques to detect credit card fraud,<sup>2</sup> as well as the use of related anomaly-detection methods to identify potential terrorist activity.<sup>3</sup> Businesses have shifted toward data-driven decision-making;<sup>4</sup> this shift is reflected in the incorporation of data-mining techniques into leading relational database management systems.<sup>5</sup> In addition to data-mining techniques, machine learning techniques<sup>6</sup> are now central to applications ranging from cars that drive themselves,<sup>7</sup> to spam filtering,<sup>8</sup> to the classification of astronomical objects.<sup>9</sup>

Although modern legal practice has adopted IT in many areas, these legal tools do not typically match the sophistication of tools found in other industries. Besides basic office software like word processors and e-mail, law firms often have comprehensive, networked document retrieval systems,<sup>10</sup> while courts and government agencies have electronic filing systems.<sup>11</sup> However, these tools lack the analytical power of IT used in other sectors of the business world. Some case management systems do include automatic text processing

---

2. See Philip K. Chan et al., *Distributed Data Mining in Credit Card Fraud Detection*, IEEE INTELLIGENT SYS., Nov./Dec. 1999, at 67.

3. See JEFFREY W. SEIFERT, DATA MINING: AN OVERVIEW (Cong. Research Serv., CRS Report for Congress Order Code RL31798, Dec. 16, 2004), available at <http://www.fas.org/irp/crs/RL31798.pdf>.

4. See generally IAN AYRES, SUPER CRUNCHERS (2007) (discussing examples of data-driven decision-making by modern businesses).

5. See, e.g., Oracle, Oracle Data Mining, <http://www.oracle.com/technology/products/bi/odm/index.html> (last visited May 12, 2008).

6. Machine learning (or statistical learning) is a general term describing a variety of techniques for automatically finding patterns in data. See TREVOR HASTIE ET AL., THE ELEMENTS OF STATISTICAL LEARNING 1–2 (2001); see also *infra* Part IV.C.

7. The 2005 DARPA Grand Challenge, a competition sponsored by the U.S. Government, was a race of autonomous land vehicles across the Mojave Desert. The winning team used machine learning to train its vehicle's driving algorithms. See Stanford Racing, <http://cs.stanford.edu/group/roadrunner/old/index.html> (last visited May 12, 2008).

8. See Mehran Sahami et al., *A Bayesian Approach to Filtering Junk E-mail*, in LEARNING FOR TEXT CATEGORIZATION: PAPERS FROM THE 1998 WORKSHOP 55 (AAAI Technical Rep. WS-98-05, 1998). In addition to Bayesian filtering, well-known spam filtering programs have used other machine learning techniques, such as genetic algorithms and feed-forward neural networks. See HowScoresAreAssigned — Spmassassin Wiki, <http://wiki.apache.org/spamassassin/HowScoresAreAssigned> (last visited May 12, 2008).

9. See Peter Cheeseman et al., The AutoClass Project, <http://ic.arc.nasa.gov/ic/projects/bayes-group/autoclass/> (last visited May 12, 2008).

10. See, e.g., Greta Rusanow, *Global Law Firm Knowledge Management Survey 2006*, LLRX.COM, Nov. 3, 2006, <http://www.llrx.com/features/kmsurvey2006.htm>.

11. See, e.g., Admin. Office of the U.S. Courts, What is PACER?, <http://pacer.psc.uscourts.gov/pacerdesc.html> (last visited May 12, 2008) (describing the electronic public access system for U.S. federal court dockets); U.S. Securities & Exchange Comm'n, SEC Filings & Forms (EDGAR), <http://www.sec.gov/edgar.shtml> (last visited May 12, 2008) (describing SEC's electronic filing system in detail).

and classification systems, however, legal professionals have not yet widely adopted these systems.<sup>12</sup> Given the extensive adoption of IT by other industries, it appears that modern legal practice has somewhat lagged behind.

There are strong incentives for legal practitioners to break this trend. Attorneys are compelled to sift through an ever-growing volume of information; the relatively underdeveloped use of IT in legal practice has left room for significant efficiency gains by eliminating repetition and wasted human resources. However, currently several barriers stand in the way of such progress. Skepticism abounds about the efficacy of artificial intelligence applications, and many technical challenges to implementation remain. Additionally, cultural resistance by the bar and legal restrictions on who may practice law are slowing adoption.

To emphasize the need for further technological development in the legal profession, this Note surveys recent developments in IT that have the potential to transform the practice of law. Part II describes some of the incentives for practitioners to adopt information technologies. Part III examines current uses of technology in the legal profession, including some recent developments of more sophisticated software. Part IV explores promising avenues of research into techniques for modeling, interpreting, and organizing information. Part V considers some of the more immediate barriers to integration of new technologies into the practice of law. Part VI concludes.

## II. INCENTIVES FOR BETTER INTEGRATION OF INFORMATION TECHNOLOGY AND LAW

Legal professionals have two primary motivations for integrating new information technologies into the practice of law. First, the volume and diversity of data that attorneys must analyze in the course of their work have exploded. Second, the efficiency gains in other industries highlight the cost savings that can be achieved by adopting more sophisticated technology.<sup>13</sup>

Legal information takes a great variety of forms. Familiar examples from litigation practice include judicial opinions, court orders, dockets, briefs, transcripts, jury instructions, and verdict statistics. There is also an enormous, but less public, body of transactional legal

---

12. See discussion *infra* Part III.

13. See generally AYRES, *supra* note 4. One measure of the perceived value of technology in other fields is the existence of well-developed academic programs, as found in computational biology and computational finance. See, e.g., Yale University, Yale Computational Biology and Bioinformatics, <http://cbb.yale.edu> (last visited May 12, 2008); Purdue University, Computational Finance at Purdue, [http://www.stat.purdue.edu/purdue\\_comp\\_finance](http://www.stat.purdue.edu/purdue_comp_finance) (last visited May 12, 2008). There is not yet any comparable programs in law.

materials — such as contracts and licenses — that shape commercial practice, even if such documents are never used in court. In addition to the core materials that would universally be considered “legal” in nature, there are many types of documents that are highly relevant to legal practice. For example, medical textbooks or expert witness reports may be relevant in personal injury cases, while purchase receipts and spreadsheets may be relevant in tax refund suits. The breadth of information types means that legal software must concern itself with written language from a diversity of sources.

Lawyers need a means for dealing with the increasing bulk of legal data. In common law jurisdictions, the body of case law expands each year: a large portion of new case law does not overrule old law, but instead refines or adapts old law to new circumstances.<sup>14</sup> Moreover, lawyers now refer to more kinds of documents in conducting their research: whereas in the print era research was largely confined to appellate cases bound in official reporters, now legal data services provide online access to “unpublished” appellate cases, lower court orders, briefs, and extra-jurisdictional materials.<sup>15</sup> In some circumstances, changes in the law itself have increased the amount of information lawyers and their clients must process. For example, in 2006, the U.S. Supreme Court approved amendments to the Federal Rules of Civil Procedure that required disclosure of a broad class of electronically stored information during litigation.<sup>16</sup> This change has increased the volume of information available during discovery beyond the high levels that already result from the United States’ liberal discovery rules.<sup>17</sup> Nor is the increase limited to litigation: the Sarbanes-Oxley Act of 2002<sup>18</sup> tightened the restrictions on the types of documents corporations must retain.<sup>19</sup> Such records are kept in anticipation of

---

14. See, e.g., EDWARD H. LEVI, AN INTRODUCTION TO LEGAL REASONING 1–8 (1949) (describing how the common law builds on itself as a process of continuous refinement). Civil law jurisdictions use cases as well, albeit with less precedential force. See Kevin D. Ashley, Case-Based Models of Legal Reasoning in a Civil Law Context (Feb. 2004), <http://www.lrdc.pitt.edu/Ashley/Mexico.pdf>.

15. See, e.g., LexisNexis, Searchable Directory of Online Sources, <http://w3.nexis.com/sources> (last visited May 12, 2008) (listing the types of publications available to search on LexisNexis).

16. See Amendments to Fed. Rules of Civil Procedure, 547 U.S. 1233, 1241 (2006).

17. See Eric Sinrod, *E-Discovery: The Times, They Are a Changing*, FINDLAW, Aug. 7, 2006, <http://technology.findlaw.com/articles/00006/010189.html>; see also Daniel Fisher, *The Data Explosion*, FORBES, Oct. 1, 2007, at 72; see generally Henry S. Noyes, *Good Cause Is Bad Medicine for the New E-Discovery Rules*, 21 HARV. J.L. & TECH. 49 (2007) (arguing that the 2006 amendments have done little to contain discovery).

18. See Pub. L. No. 107-204, 116 Stat. 745 (codified at 18 U.S.C. § 1520 (Supp. V 2005)).

19. *Id.* § 802(a), 116 Stat. at 800 (authorizing the Securities and Exchange Commission to promulgate regulations relating to document retention, and requiring accountants to retain audit and review workpapers of securities issuers for five years); see also Michele C.S. Lange, *New Act Has Major Impact on Electronic Evidence*, NAT’L L.J., Nov. 4, 2002, at C8,

future use in regulatory compliance or litigation.<sup>20</sup> This explosion in the number and type of documents with which attorneys must concern themselves is an open invitation for technological innovation.

One of the primary problems for firms affected by the document explosion is processing and understanding the growing volume of information they manage. Because IT solutions are not available, firms increasingly rely on contract attorneys to assist with document-intensive matters, despite the considerable cost.<sup>21</sup> In time, new computational techniques may supplement or supplant this practice.

Although contemporary legal practice incurs significant costs because of repetitive inefficiencies, new technologies can potentially produce considerable savings. Some of the documents that lawyers currently handle are already structured in limited ways that are amenable to computer-reading — for instance, the federal district courts of the Northern District of California require motions to contain the case number, date and time of a hearing at particular locations in the document.<sup>22</sup> For the most part, however, legal data is far less structured than the tabular data in a relational database or a spreadsheet.<sup>23</sup> Newer, moderately sophisticated technologies like document assembly — the computer-assisted production of documents like contracts — can reduce the number of attorneys necessary to draft a given document.<sup>24</sup> Researchers who develop new information technologies describe a future with intelligent computerized legal assistants that can scour databases and outline arguments in place of low-level associates,<sup>25</sup> as well as sophisticated software agents that can negotiate contracts without the direct involvement of attorneys.<sup>26</sup> The realization of such cost saving techniques could dramatically alter the land-

---

available at <http://www.law.com/jsp/article.jsp?id=1039054510969> (describing the new retention requirements and their impact on electronic data management).

20. See Lange, *supra* note 19.

21. See Leigh Jones, *More Firms Using Temp Attorneys*, NAT'L L.J., Oct. 10, 2005, at 1, available at <http://www.law.com/jsp/ihc/PubArticleIHC.jsp?id=1128947761813>.

22. See N.D. CAL. CIV. R. 7-2(b) (describing the form of motions submitted to the court), available at <http://www.cand.uscourts.gov/CAND/LocalRul.nsf>.

23. Relational databases are similar to spreadsheets but have greater structural constraints while allowing for more sophisticated extraction of data. See PHILIP GREENSPUN, PHILIP AND ALEX'S GUIDE TO WEB PUBLISHING 333–44 (1999).

24. See RICHARD SUSSKIND, THE FUTURE OF LAW 215–17 (1996) (arguing that systematizing the legal knowledge and expertise necessary to draft a contract has the potential to fundamentally reshape the legal process); Darryl R. Mountain, *Disrupting Conventional Law Firm Business Models Using Document Assembly*, 15 INT'L J.L. & INFO. TECH. 170 (2007) (expanding Susskind's argument by analyzing business models).

25. Kevin Ashley, *Case-Based Reasoning*, in INFORMATION TECHNOLOGY AND LAWYERS, *supra* note 1, at 23, 23–24.

26. Edwina L. Rissland, Kevin D. Ashley & R.P. Loui, *AI and Law: A Fruitful Synergy*, 150 ARTIFICIAL INTELLIGENCE 1, 15 (2003) (speculating on future legal applications of artificial intelligence).

scape of legal practice by eliminating much of the wasteful repetition that many practitioners have observed.<sup>27</sup>

Thus, both the growth in the volume of documents attorneys must handle and potential cost-savings from efficiency gains offer significant motivations for legal practitioners to adopt better information technologies. The next Part discusses the current state of the legal profession's adoption of technologies in light of these incentives.

### III. THE CURRENT STATE OF INFORMATION TECHNOLOGY IN LEGAL PRACTICE

In addition to basic office automation software, lawyers already use a wide range of computational tools of varying sophistication. These tools include databases of legal materials, software for document assembly, and software for litigation support.

A typical law firm uses software similar to that found in most organizations: word processors, e-mail systems, file servers, and the like. Since these tools can accommodate any sort of textual data, they are useful for working with even the most heterogeneous collections of legal materials. Also in widespread use, however, are a number of more law-specific technologies. Legal calendar software extends the basic project management software by including legal timing rules.<sup>28</sup> Billing software tracks billable hours and integrates billing information into accounting and financial software packages.<sup>29</sup> Conflict management software tracks potential conflicts of interest among a firm's clients and potential clients.<sup>30</sup>

Electronic data services like LexisNexis and Westlaw have become firmly entrenched as legal research tools.<sup>31</sup> The primary appeal of these services is likely their comprehensive coverage and compilation of legal documents, rather than any particular technological feature. The typical interface is a relatively straightforward keyword search, which allows users to search either the primary text of the various documents or a selection of fields, such as author, title, and other surface features.<sup>32</sup> In addition to keyword search, these services

---

27. See, e.g., Lauritsen, *supra* note 1, at 175.

28. See, e.g., AbacusLaw, Legal Calendar Software — Popular Features, <http://www.abacuslaw.com/products/abacus-law-software-features.html> (last visited May 12, 2008) (describing a software calendar that automatically schedules according to the deadlines in a particular jurisdiction).

29. See, e.g., *id.*

30. See, e.g., *id.*

31. See, e.g., Linda M. Furllet & Craig B. Simonsen, *Searching in Legal Databases on the Internet*, in *COMPUTER-AIDED LEGAL RESEARCH (CALR) ON THE INTERNET* 79, 79–80 (Craig B. Simonsen & Christian R. Anderson eds., 2006).

32. Efficient keyword search of large databases is not a trivial task. Google based its successful search engine on an algorithm for ranking search results to favor websites linked to by other highly ranked websites. See Lawrence Page et al., *The PageRank Citation Ranking*:

offer more sophisticated features for accessing and understanding legal documents. Online documents may link to the documents they cite, and may also provide indices of the documents that cite to them.<sup>33</sup> These reverse-citation indices come with indications — prepared by humans — of whether the case or statute in question is cited favorably, overruled, or distinguished on some point of law or fact.<sup>34</sup> Humans also prepare summaries of cases, subject classifications, and “headnotes” (concise statements of legal issues and holdings in a court opinion).<sup>35</sup> Although generating such features is laborious, the features do encode a substantial amount of structured, semantic content into electronically-searchable material, which can be helpful not only for keyword searches but also for enabling some of the more advanced techniques discussed in Part IV.

Legal data services have developed some additional tools that generate useful information in a more automated way. West’s Case Evaluator system uses a forms-based interface to collect information about a case and automatically assembles reports that include relevant case citations, verdict trends for similar cases in the jurisdiction, and excerpts from relevant expert materials.<sup>36</sup> For transactional practice, West’s Deal Proof tracks certain key phrases in contracts — in particular defined terms and repeated, legally-significant phrases — and includes tools to ensure that those definitions and phrases are used appropriately and consistently throughout a document.<sup>37</sup> West’s software also uses automatic text classification to identify and recommend documents likely to be related to other documents that the user has already located.<sup>38</sup>

---

Bringing Order to the Web (1998), <http://dbpubs.stanford.edu/pub/1999-66>. The legal data services use a different strategy to rank results, making use of Bayes’ Theorem to identify results most likely to reflect what the user sought based on the search terms. *See* Howard Turtle, *Text Retrieval in the Legal World*, 3 *ARTIFICIAL INTELLIGENCE & L.* 5, 31–34 (1995). It is also possible to submit queries in natural language. This feature is mostly offered as a convenience to users and does not allow users to search in ways that are fundamentally different from those of a more structured query. *See* West, *Natural Language Searches*, <http://lawschool.westlaw.com/shared/marketinfodisplay.asp?code=RE&id=151&subpage=&rtcode=re&rtid=159> (last visited May 12, 2008).

33. *See, e.g.*, LexisNexis, Shepard’s Citations Service, <http://law.lexisnexis.com/shepards> (last visited May 12, 2008).

34. *See, e.g., id.*

35. *See, e.g., id.*

36. *See* WEST, MATERIAL NO. 40687294, WESTLAW QUICK REFERENCE GUIDE: CASE EVALUATOR 1 (Feb. 2008), <http://west.thomson.com/documentation/westlaw/wlawdoc/wlres/caseeval.pdf>.

37. *See* WEST, ITEM NO. 1-6136-5, DEAL PROOF QUICK REFERENCE GUIDE: DEAL PROOF TUTORIAL 3 (Oct. 2007), <http://west.thomson.com/documentation/westlaw/wlawdoc/wlres/wkmdptut.pdf>.

38. *See* Peter Jackson, *Generating Value from Textual Discovery*, in *COMPUTATIONAL SCIENCE — ICCS 2007: 7TH INTERNATIONAL CONFERENCE PROCEEDINGS, PART III*, at 746, 749–51 (Lecture Notes in Computer Sci., Vol. No. 4489, 2007) (describing the mechanics of West’s ResultsPlus product).

Among the most sophisticated computational tools used by law firms are document assembly tools like HotDocs and DealBuilder.<sup>39</sup> Document assembly tools prompt the lawyer to enter information about the facts and issues involved in a matter.<sup>40</sup> Unlike West's Case Evaluator, which is also driven by input forms,<sup>41</sup> these assembly tools use document templates to construct legal documents like complex contractual agreements, which can then be reviewed and edited by human lawyers.<sup>42</sup> The Dutch and Flemish governments have already used similar legal drafting systems to assist lawmakers in drafting of statutes.<sup>43</sup> In large firms, the document assembly systems often rely on in-house templates and model documents, and a great deal of effort goes into customizing the documents for general use.<sup>44</sup> Perhaps indicating the direction of future developments, a British company offers drafting systems with built-in templates for a few well-defined areas of law.<sup>45</sup>

A number of commercial services attempt to streamline the discovery process in large-scale litigation.<sup>46</sup> In some cases, the express goal is to replace the traditional discovery method, which consists of teams of junior attorneys working within a law firm.<sup>47</sup> Practitioners have had some success with these services, reporting only small numbers of false positives and false negatives from searches through large sets of discoverable documents.<sup>48</sup> Such litigation support systems use some of the same machine learning and legal ontology techniques

---

39. See Lauritsen, *supra* note 1, at 168; see also, e.g., Bus. Integrity Ltd., Law Firms Applications, <http://www.business-integrity.com/applications-law-firms.html> (last visited May 12, 2008).

40. See, e.g., Bus. Integrity Ltd., *supra* note 39.

41. See WEST, *supra* note 36, at 3.

42. See Bus. Integrity Ltd., *supra* note 39.

43. Marie-Francine Moens, *Improving Access to Legal Information: How Drafting Systems Help*, in INFORMATION TECHNOLOGY AND LAWYERS, *supra* note 1, at 119, 122–24.

44. See, e.g., HotDocs, HotDocs Products, <http://hotdocs.com/products> (last visited May 12, 2008) (describing a software application that converts word processor files into interactive templates).

45. See Practical Law Co., PLC — FastDraft, <http://www.practicallaw.com/1-203-1674> (last visited May 12, 2008). Richard Susskind has suggested that this software may be a harbinger of a transition from individualized legal counseling to an off-the-shelf product, at least for relatively straightforward matters. See Richard Susskind, *Podcasts for Desperate and Diligent Students*, TIMES (London), Sept. 5, 2006, at Law 7, available at [http://business.timesonline.co.uk/tol/business/law/columnists/richard\\_susskind/article626363.ece](http://business.timesonline.co.uk/tol/business/law/columnists/richard_susskind/article626363.ece).

46. See, e.g., Discovery Mining, Discovery Mining Automated Discovery Platform, <http://www.discoverymining.com/discovery/index.html> (last visited May 12, 2008); H5, Services, <http://www.h5technologies.com/services/index.html> (last visited May 12, 2008) (describing services for outsourcing aspects of litigation discovery and regulatory compliance by using machine learning methods for document classification and information retrieval).

47. See, e.g., H5, DIFFERENTIATION 1 (2006), <http://www.h5technologies.com/pdf/differentiation.pdf>.

48. See Anne Kershaw, *Talking Tech: Automated Document Review Proves Its Reliability*, DIGITAL DISCOVERY & E-EVIDENCE, Nov. 2005, at 10, 12.

discussed in Part IV.B and IV.C.<sup>49</sup> A typical application is to search through thousands of documents and to flag those sufficiently similar to a model document.<sup>50</sup>

As these examples suggest, there is a great deal of technology already in use in legal practice, some of it surprisingly forward-looking. Even the most sophisticated systems in use, however, depend on extensive human intervention to achieve useful results. Moreover, the systems generally restrict themselves to a relatively superficial analysis of the underlying legal texts and lack schemes for detailed knowledge representation or automatic processing of legal texts based on semantic content. The research systems described in Part IV aim to fill some of these gaps.

#### IV. THE DIRECTION OF LEGAL INFORMATICS: CURRENT RESEARCH

This Part describes three major threads of research in IT and law: models of legal argumentation with explicit mechanisms for representing knowledge, applications of machine learning techniques to legal data, and ways of organizing and distributing legal information.

##### *A. Advances in Argumentation Models and Outcome Prediction*

Computer tools for automating aspects of legal practice could take a number of forms. These forms might include more powerful document search mechanisms that consider the semantics of the documents<sup>51</sup> as well as the plain text, software assistants that produce legal arguments corresponding to one litigant's perspective, and predictive tools that assess the probabilities of several possible outcomes given information about a case.<sup>52</sup> Much of the recent research in artificial intelligence and law has focused on techniques that could be useful for all of these purposes.<sup>53</sup> Such techniques are typically described as "adversarial case-based reasoning systems"<sup>54</sup> or argumentation systems.<sup>55</sup>

---

49. See *infra* Part IV.B–C; Kershaw, *supra* note 48, at 11.

50. See Kershaw, *supra* note 48, at 11.

51. An example of a semantic search involves limiting results to cases in which the plaintiff prevailed on the issue of interest. There is no easy way to do this unless either the search tool can infer who is the winning party from context, see *infra* Part IV.B (discussing machine learning and text mining), or issues are pre-marked with a tag indicating which party won, see *infra* Part IV.C (discussing semantic tagging).

52. See Ashley, *supra* note 25, at 26.

53. See *id.* at 23–25.

54. *Id.* at 25 (describing systems that consider "cases . . . to justify how a problem situation should be decided").

55. See ANNE VON DER LIETH GARDNER, AN ARTIFICIAL INTELLIGENCE APPROACH TO LEGAL REASONING 38–41 (1987); Trevor Bench-Capon & Henry Prakken, *Argumentation*,

Prior to the development of case-based reasoning systems, Anne von der Lieth Gardner developed a rule-based system for applying contract law to fact patterns.<sup>56</sup> Gardner extracted substantive rules from the Restatement of Contracts and supplemented them with interpretive rules designed to allow the Restatement's rules to be applied to facts.<sup>57</sup> The system classified cases as "easy" if the rules yielded a unique result or "hard" if they did not; for the easy cases it provided outcome predictions.<sup>58</sup> Gardner's system resembles the rule-based expert systems that were popular in the 1970s and 1980s, both in artificial intelligence research and in industry.<sup>59</sup> Although similar systems have proven moderately useful for assisting human decision-makers,<sup>60</sup> they suffer from certain deficiencies. First, in order to create the legal precedent, rule-based systems require a laborious intermediate step in which a human assembles a coherent set of rules.<sup>61</sup> Second, although the systems can identify issues and reach conclusions, they cannot generate arguments supporting a particular litigant's position.<sup>62</sup> Third, rule-based systems are not able to weigh factors or apply multi-part balancing tests.<sup>63</sup> Fourth, the systems have difficulty evaluating legal arguments that depend on nuanced factual or procedural contexts.<sup>64</sup>

Recent case-based systems have achieved some success in mitigating these problems.<sup>65</sup> The fundamental principle on which the case-based systems operate is that "[a] particular party in a given scenario should win a claim or an issue because a similarly situated party won such a claim or issue in a particular case whose facts are relevantly similar and where the same or similar law applied."<sup>66</sup> This principle parallels the rationale behind the treatment of precedent in the common law tradition. Two systems designed to implement this principle in research and teaching contexts are Hypo and CATO.

Developed by Professor Kevin Ashley, Hypo uses precedents to construct arguments for one side in a trade secret dispute, and then

---

in INFORMATION TECHNOLOGY AND LAWYERS, *supra* note 1, at 61, 62–63 (describing systems that use logical methods to model persuasive arguments).

56. See Bench-Capon & Prakken, *supra* note 55, at 63.

57. See *id.*

58. See *id.* at 63–64; GARDNER, *supra* note 55, at 38–41.

59. See, e.g., STUART J. RUSSELL & PETER NORVIG, ARTIFICIAL INTELLIGENCE: A MODERN APPROACH 22–24 (2d ed. 2003) (describing the rule-based expert systems of the 1970s and 1980s).

60. See, e.g., Bench-Capon & Prakken, *supra* note 55, at 64–65 (discussing a rule-based system designed to help officers decide whether to approve environmental permits).

61. See GARDNER, *supra* note 55, at 85–162 (describing the process of representing problems and defining legal rules).

62. See Bench-Capon & Prakken, *supra* note 55, at 65.

63. See *id.*

64. See *id.* An example of such an argument is a fact-specific inquiry into the legality of a religious display on public property. See, e.g., *County of Allegheny v. ACLU*, 492 U.S. 573 (1989).

65. See Bench-Capon & Prakken, *supra* note 55, at 65–66.

66. Ashley, *supra* note 25, at 35.

constructs counter-arguments for the other side by citing alternative precedents and distinguishing the first side's cases.<sup>67</sup> Cases are represented as having a particular position with respect to a number of dimensions that might affect the plausibility of an argument.<sup>68</sup> Because Hypo deals with trade secret law, relevant dimensions include the extent to which a plaintiff took security precautions to protect its secrets, and whether the secrets were disclosed to the defendant.<sup>69</sup>

Building on the same principles underlying Hypo, Vincent Aleven designed the CATO system as a teaching aid to help law students learn to argue from precedent.<sup>70</sup> Although CATO uses only binary-valued factors, it organizes these factors in a case-specific hierarchy that "provides legal reasons why trade secret factors matter in terms of more abstract [f]actors."<sup>71</sup> Vincent Aleven designed CATO to account for background knowledge in a context-sensitive manner, such that the significance of a particular case depends on the purpose of the argument in which the case is used.<sup>72</sup> CATO produces the text of an argument for one side of an issue in plain English, and includes a graphical representation of the argument structure it created.<sup>73</sup>

After Hypo and CATO, newer developments have included more detailed representation of arguments with precedential cases, increased emphasis on outcome prediction and argument generation, and improved predictive accuracy in simulations. For example, the GREBE system contains representations of the semantic structure of the rationales of the cases in its database.<sup>74</sup> These representations allow the system to rearrange arguments that appear in precedential cases by extracting sub-rules and drawing structural analogies.<sup>75</sup> No-

---

67. *See id.* at 39.

Basically Hypo takes as its input a fact situation describing a trade secrets dispute . . . . The outputs show the best precedents to cite for each side on the claim, how those cases may be cited in legal points, how an opponent would respond to each point, and how a point or response could be strengthened by the addition or subtraction of crucial facts.

KEVIN D. ASHLEY, *MODELING LEGAL ARGUMENT* 36 (1990).

68. *See Ashley, supra* note 25, at 37–38.

69. *See id.*

70. *See Vincent Aleven, Using Background Knowledge in Case-Based Legal Reasoning: A Computational Model and an Intelligent Learning Environment*, 150 *ARTIFICIAL INTELLIGENCE* 183, 184–90 (2003).

71. *See Ashley, supra* note 25, at 39.

72. Aleven, *supra* note 70, at 184.

73. *See id.* at 196 (showing sample output); *id.* at 221 (showing the graphical interface).

74. *See L. Karl Branting, A Reduction-Graph Model of Precedent in Legal Analysis*, 150 *ARTIFICIAL INTELLIGENCE* 59, 64–74 (2003); *see also Ashley, supra* note 25, at 43–46.

75. *See Branting, supra* note 74, at 76–77.

tably, GREBE produces formatted sentences, rather than the condensed shorthand of some of the other research systems.<sup>76</sup>

Several of these systems have achieved relatively high prediction accuracy. Similar to Hypo, the Issue-Based Prediction (“IBP”) program models trade secret misappropriation arguments and predicts the winner of a dispute with 91.4% accuracy.<sup>77</sup> A group of researchers based in New Zealand has developed a multi-step method for intelligent retrieval of precedents from a database; the method was able to identify 96.3% of the precedents cited in real judicial opinions.<sup>78</sup> The majority of the precedents that the program did not successfully identify appeared in dicta or in distinguishing citations.<sup>79</sup>

Professor Ashley suggests that useful systems will likely combine many different techniques and will need the capacity to search databases for precedential authority, construct arguments, and predict outcomes.<sup>80</sup> To succeed in the legal world, such systems will also have to integrate well with existing legal data services like LexisNexis and Westlaw.<sup>81</sup> Ultimately, while IT has made considerable progress in modeling legal argument, it must still overcome additional obstacles before more practical implementations can be widely deployed.

### *B. Machine Learning and Knowledge Discovery from Databases*

“Machine learning” usually refers to a panoply of techniques taken from artificial intelligence, statistics, and other fields.<sup>82</sup> The goal is to glean non-obvious information from large data sets, where the structures are often more complex than in data sets that can be analyzed using traditional statistical regression models.<sup>83</sup> Machine learning algorithms fall into three broad classes: supervised, unsupervised, and reinforcement learning.<sup>84</sup> Supervised learning uses a “training” data set, in which certain inputs generate known outputs, and the

---

76. See, e.g., Ashley, *supra* note 25, at 46 fig.9 (showing an example of the human-readable output produced by GREBE).

77. See *id.* at 53–55; see also Kevin D. Ashley & Stefanie Brüninghaus, *Computer Models for Legal Prediction*, 46 JURIMETRICS 309, 347 (2006).

78. Yiming Zeng et al., *A Knowledge Representation Model for the Intelligent Retrieval of Legal Cases*, 15 INT’L. J.L. & INFO. TECH. 299, 314 (2006). Zeng’s method narrows the search range in three ways: by identifying key issues in the fact pattern, by considering the presence of pro-claimant or pro-respondent factors, and by weighing neutral contextual features. *Id.* at 304–05.

79. *Id.* at 314.

80. Ashley, *supra* note 25, at 23–26.

81. See *id.*; *supra* notes 31–37 and accompanying text for a discussion of existing legal data services.

82. See, e.g., RUSSELL & NORVIG, *supra* note 59, at 712–88 (describing the concept using the equivalent phrase “statistical learning”).

83. Some common machine learning algorithms include decision trees, neural networks, and support vector machines. See *id.*

84. *Id.* at 650.

learning algorithm minimizes error in output prediction.<sup>85</sup> Supervised learning on categorical data is often called classification.<sup>86</sup> In unsupervised learning algorithms, input-output pairs are unknown.<sup>87</sup> Both supervised and unsupervised learning methods have been applied to legal problems.<sup>88</sup>

Andrew Stranieri and John Zeleznikow have emphasized the important role that machine learning can play in uncovering and quantifying the “open texture” of legal rules and the discretion of individual legal decision-makers.<sup>89</sup> They claim that machine learning is especially well-suited to predicting outcomes dependent on “local stare decisis” (when the same decisions follow from similar fact patterns before the same court) or “personal stare decisis” (when the same decisions follow from similar fact patterns before the same judge), rather than on traditional stare decisis.<sup>90</sup> In particular, machine learning more effectively predicts outcomes in ordinary cases that depend on judicial discretion than in cases that do not announce broader changes in legal doctrine.<sup>91</sup>

Other applications of machine learning include extraction of legal rules from databases,<sup>92</sup> measurement of trends in the application of rules over time,<sup>93</sup> and identification of clusters of related cases or documents.<sup>94</sup> Some methods can be combined; for instance, machine learning techniques can be used to automate the data preparation for the argumentation systems discussed in Part IV.A.<sup>95</sup> Systems such as Hypo currently require a human to enumerate the factors or delineate the argument structure for all of the cases in their databases.<sup>96</sup> Thus, there are obvious scalability concerns for systems intended to be used with larger databases: the databases must be created by humans. A machine learning program that could extract factors and rules from cases in a sufficiently uniform and reliable way may be the only cost-

85. *Id.*

86. *Id.* at 653.

87. *Id.* at 650.

88. See generally ANDREW STRANIERI & JOHN ZELEZNIKOW, KNOWLEDGE DISCOVERY FROM LEGAL DATABASES (2005) (discussing numerous systems that have applied machine learning algorithms to legal problems). The third class, reinforcement learning, applies to problems in which the environment provides some feedback to the learning agent about how well it is performing. See RUSSELL & NORVIG, *supra* note 59, at 650. The author does not know of any applications of reinforcement learning methods to legal problems.

89. Andrew Stranieri & John Zeleznikow, *Knowledge Discovery from Legal Databases — Using Neural Networks and Data Mining to Build Legal Decision Support Systems*, in INFORMATION TECHNOLOGY AND LAWYERS, *supra* note 1, at 81, 82.

90. *Id.* at 87–88.

91. See *id.* at 84–87; see also STRANIERI & ZELEZNIKOW, *supra* note 88, at 214–16.

92. STRANIERI & ZELEZNIKOW, *supra* note 88, at 95–96.

93. See Stranieri & Zeleznikow, *supra* note 89, at 112–13.

94. See *id.*, at 111–12; Turtle, *supra* note 32, at 26 (describing cluster-based models for document retrieval).

95. See Ashley, *supra* note 25, at 56–58.

96. See ASHLEY, *supra* note 67, at 25–34.

effective way to deploy argumentation systems using large databases.<sup>97</sup>

Machine learning algorithms have also been used for more direct decision-making support. The “Judges on Wheels” program in Brazil sends police, an insurance adjustor, and a judge to the scene of traffic collisions.<sup>98</sup> The judge is advised by a computer program that has been trained using past decisions of other judges.<sup>99</sup> Although the judges are not obliged to accept the program’s recommendation, they do so 68% of the time.<sup>100</sup> A drawback of the system is that the machine learning methods used frustrate extraction of the rationale that the program used to reach its recommendation: usually, the most one can say is that the decision accords with the decisions used to train the software.<sup>101</sup>

Machine learning algorithms have already demonstrated some capacity to assist legal decision-makers. Further development of these methods has the potential to reduce inefficiencies and bolster the productivity of legal practitioners.

### C. Accessible, Structured Knowledge

For computerized applications, the unstructured character of most legal data presents a technical problem. Typically, unstructured documents require substantial pre-processing before they can be ana-

---

97. An early attempt to use automated information extraction methods to prime a case-based reasoning system is described in Stefanie Brüninghaus & Kevin D. Ashley, *Improving the Representation of Legal Case Texts with Information Extraction Methods*, 2001 PROC. 8TH INT’L CONF. ON ARTIFICIAL INTELLIGENCE & L. 42, 46–47. See generally Rosina O. Weber, Kevin D. Ashley & Stefanie Brüninghaus, *Textual Case-Based Reasoning*, 20 KNOWLEDGE ENG’G REV. 255 (2006) (discussing the state of the art of information extraction from textual sources). The amount of work to create a case-based reasoning system could be substantial. The CYC project was started by AI researcher Douglas Lenat in 1984 to develop a large database of commonsense background facts to serve as a foundation for future AI work. See Douglas B. Lenat, *CYC: A Large-Scale Investment in Knowledge Infrastructure*, 38 COMM. ACM 32, 33 (1995). Although the project now gets most of its new facts by data-mining the Web, for most of its two-decade history it employed people to enter millions of facts. See Cynthia Matuszek et al., *Searching for Common Sense: Populating Cyc from the Web*, in 2005 PROC. TWENTIETH NAT’L CONF. ON ARTIFICIAL INTELLIGENCE & SEVENTEENTH INNOVATIVE APPLICATIONS OF ARTIFICIAL INTELLIGENCE CONF. 1430, 1430. The result is a system that answers natural-language queries about all sorts of information available on the Web, along with explanations of how the answers were derived.

98. See Stranieri & Zeleznikow, *supra* note 89, at 102.

99. See *id.*

100. *Id.*

101. See *id.*

lyzed by computers.<sup>102</sup> The emerging “Semantic Web” may help to alleviate the need for some of this pre-processing.<sup>103</sup>

The Semantic Web is a general framework wherein syntax is designed to model semantics more closely than conventional online markup languages like HTML currently allow.<sup>104</sup> Such modeling better enables machine-readers to extract the information about which human users care.<sup>105</sup> The framework comprises three main parts: first, a formal language that allows document creators to specify relationships among the concepts they employ; second, a high-level “ontology” that specifies the rules governing valid manipulations of the relationships within the relevant domain of knowledge; and third, a flexible format for storing data.<sup>106</sup>

Placing legal information — e.g., statutes, regulations, and judicial opinions — into the Semantic Web will enable search tools and decision support systems to operate on uniformly structured data, without relying on more uncertain methods for extracting information from plain text.<sup>107</sup> Machine learning methods will be able to identify rules and patterns more accurately in such a data set.<sup>108</sup> The Semantic Web approach does have disadvantages: the development of suitable ontologies and the formatting of appropriately structured documents can prove particularly difficult.<sup>109</sup> Nevertheless, the potential rewards are great enough that there are a number of projects devoted to encoding whole areas of law in the structured language of the Semantic Web.<sup>110</sup>

In the United States, the Semantic Web has been employed as part of a broader movement to provide free access to legal informa-

102. See STRANIERI & ZELEZNIKOW, *supra* note 88, at 147–69 (describing the problem of extracting information from unstructured text); *see also id.* at 47–58 (describing techniques for dealing with missing and inconsistent data).

103. See V. Richard Benjamins et al., *Law and the Semantic Web, an Introduction*, in LAW AND THE SEMANTIC WEB 1, 2 (V. Richard Benjamins et al. eds., 2005).

104. See Tim Berners-Lee et al., *The Semantic Web*, SCI. AM., May 2001, at 34, 36–37.

105. *See id.*

106. *See id.* at 38–42 (describing a system that uses the Resource Description Framework (“RDF”) to define relationships among concepts in conjunction with the eXtensible Markup Language (“XML”) to structure data). Berners-Lee et al. provide an example of a simple, hypothetical Semantic Web agent that runs on a handheld computer and is able to schedule a medical appointment subject to constraints on location, time, other appointments in the user’s schedule, et cetera. The agent presumably makes use of a knowledge representation scheme capable of representing the relevant factors, and an inference mechanism to apply the constraints in the problem at hand. *See id.* at 36.

107. *See* Benjamins et al., *supra* note 103, at 4–5.

108. *See* STRANIERI & ZELEZNIKOW, *supra* note 88, at 204–09.

109. *See* Benjamins et al., *supra* note 103, at 9–10 (describing the greater complexity of developing legal ontologies than of developing medical or engineering ontologies); *see also* STRANIERI & ZELEZNIKOW, *supra* note 88, at 147–69 (describing the problem of extracting information from unstructured text).

110. *See, e.g.,* Estrella, EstrellaWiki, <http://www.estrellaproject.org/index.php?title=Estrella&oldid=2385> (as of May 12, 2008, 12:00 GMT) (describing a European project to encode legislation in machine-readable form, starting with European tax legislation).

tion. The public interest organizations Creative Commons and Public.Resource.Org have recently made a large fraction of U.S. case law freely available in a structured form.<sup>111</sup> Such projects are a new source of competition for established legal data services such as Westlaw and LexisNexis. Moreover, because these projects emphasize providing the public with broader access to legal sources, they may eventually compete with lawyers as well.<sup>112</sup>

In this way, research into better ways to structure legal information, better methods to extract latent patterns, and better systems for machine analysis offers the legal industry a significant opportunity. If realized, these nascent technologies may enable attorneys to reap substantial efficiency gains by eliminating much of the redundancy in contemporary legal work.

#### V. INFORMATION TECHNOLOGY AND THE LEGAL PROFESSION: BARRIERS TO PROGRESS

Despite growing pressure to find new ways to manage information,<sup>113</sup> advocates of more widespread adoption of sophisticated IT in law face a number of potential barriers. Because the technologies applied to law are no different from other artificial intelligence technologies, the generalized criticisms that technology cannot replicate the human mind apply and will persist.<sup>114</sup> Cyrus Tata has argued that legal decision-making is inherently holistic and context-dependent, and has suggested that modeling such decision-making requires more than simple legal rules and formal logic.<sup>115</sup> However, the logical style of most legal writing and the predictive accuracy of some prototype systems seem to undercut this view.<sup>116</sup>

Other objections are based on the misconceived notion that artificial intelligence can do nothing but derive consequences from posited

---

111. See Press Release, Creative Commons & Public.Resource.Org, 1.8 Million Pages of U.S. Case Law Available Now for Developers: No Restrictions on Reuse (Feb. 11, 2008), available at [http://bulk.resource.org/courts.gov/0\\_Press\\_20080211.pdf](http://bulk.resource.org/courts.gov/0_Press_20080211.pdf) ("Practical access for all Americans to legal cases and material is essential to the rule of law. The Legal Commons is an important step in reducing the barriers to effective representation of average citizens and public interest advocates."); see also AltLaw, About AltLaw, <http://www.altlaw.org/v1/about> (last visited May 12, 2008) (describing a related project to make recent U.S. Supreme Court and federal appellate case law freely available online).

112. See *infra* note 125 and accompanying text for a discussion of concerns relating to the unauthorized practice of law.

113. See *supra* Part II.

114. See generally John R. Searle, *Chinese Room Argument*, in THE MIT ENCYCLOPEDIA OF THE COGNITIVE SCIENCES 115 (Robert A. Wilson & Frank C. Keil eds., 1999) (arguing that computers cannot have minds in the sense that people do).

115. See Cyrus Tata, *The Application of Judicial Intelligence and 'Rules' to Systems Supporting Discretionary Judicial Decision-Making*, 6 ARTIFICIAL INTELLIGENCE & L. 203, 223-25 (1998).

116. See *supra* Part IV.A.

sets of rules using deductive logic.<sup>117</sup> Although such programs exist, other systems account for factual uncertainty using Bayesian statistics,<sup>118</sup> legal indeterminacy using fuzzy logic,<sup>119</sup> and other forms of open texture including categorical uncertainty and vagueness of terms.<sup>120</sup> Moreover, argumentation systems can find support for a particular position, even when there is no unique, deterministic outcome.<sup>121</sup>

Objections based on the practical efficacy of using various computational techniques in a legal context are harder to dismiss. For example, some commentators have claimed that it is an empirical fact that neural networks are ineffective when applied to certain legal problems if not combined with human-generated doctrinal rules.<sup>122</sup> Even if a method is theoretically possible, the costs of its development (i.e., the costs of adapting it to a particular legal domain and preparing the requisite databases) may be prohibitively high even when compared to the ongoing costs of human labor.<sup>123</sup> It may simply take a great deal of investment to develop new systems of the complexity necessary to be useful, even if individual components have already proven effective in isolated research settings. Further investment will also be necessary to integrate new tools into mainstream legal research systems and conventional software.

Apart from possible design challenges, legal barriers may stand in the way of a company or organization developing new legal software. Although the Internet provides easy access to data, and court documents are in the public domain in the United States, there are nevertheless copyright restrictions on compilations and databases. These restrictions inhibit easy access to complete sets of legal documents without expensive negotiations with copyright owners.<sup>124</sup> More significantly, in some jurisdictions it may constitute the practice of law without a license to provide sophisticated self-help tools like those

---

117. See JOHN HAUGELAND, *ARTIFICIAL INTELLIGENCE: THE VERY IDEA* 9–12 (1985) (disputing the notion that the fact of having been programmed imposes a theoretical limitation on artificial intelligences).

118. See Ashley, *supra* note 25, at 31–34 (describing the use of Bayesian inference networks for retrieval of legal texts).

119. See STRANIERI & ZELEZNIKOW, *supra* note 88, at 111–15.

120. See *id.* at 25–31 (discussing tasks suitable for knowledge discovery from databases).

121. See Bench-Capon & Prakken, *supra* note 55, at 66–71.

122. See Dan Hunter, *Looking for Law in All the Wrong Places: Legal Theory and Legal Neural Networks*, in *LEGAL KNOWLEDGE BASED SYSTEMS: JURIX '94*, at 55, 59–61 (H. Prakken et al. eds., 1994) (arguing that several implementations have used training sets that were too small to achieve statistically significant results and that successes were influenced by the designers' implicit adherence to rule positivism).

123. Cf. Matuszek et al., *supra* note 97 at 1430 (describing the process of representing three million facts in the CYC knowledge base over the course of twenty years).

124. Cf. *Key Publ'ns, Inc. v. Chinatown Today Pub. Enters., Inc.*, 945 F.2d 509, 512–14 (2d Cir. 1991) (holding that a yellow pages directory was entitled to copyright protection as a compilation). *But see supra* Part IV.C (describing projects to make statutes, regulations, and case law freely available online).

described in Part IV.<sup>125</sup> Unlicensed practice concerns would not, however, apply for software designed for professional consumption.

Finally, the legal community itself may resist adopting sophisticated tools. This resistance may derive from professional conservatism or from a general resistance to lay involvement in the legal process.<sup>126</sup> Lawyers may feel that advanced IT techniques do not always address real problems. They may also find the tools too cumbersome to use or imperfectly integrated with existing IT infrastructure.<sup>127</sup> Additionally, economic forces within the legal profession may make it difficult for new technologies to gain acceptance in traditional firms because there is little incentive to increase efficiency.<sup>128</sup> Many new tools are designed to assist with the discovery process, however many firms pass on costs of discovery and legal research to their clients.<sup>129</sup> Furthermore, currently many firms appear to insulate themselves from price competition such that they have little incentive to reduce costs to clients.<sup>130</sup> Moreover, law firms usually have limited ability to raise outside capital and are hesitant to make significant capital investments in systemizing repetitive tasks.<sup>131</sup> In the end, it may be client demands that drive firms to adopt new technologies, rather than initiatives from within firms themselves.<sup>132</sup>

---

125. See, e.g., *In re Reynoso*, 477 F.3d 1117 (9th Cir. 2007) (finding that the operation of a website which, in exchange for a fee, generated the forms for bankruptcy filings based upon user inputs, constituted unauthorized the practice of law). *But cf.* H&R Block TaxCut, Terms of Service Agreement, <http://www.taxcut.com/universal/legal.html> (last visited May 12, 2008) (disclaiming that any communications provide legal advice). Although the IRS permits the use of tax software, users of such software may be held responsible for errors resulting from bad inputs. See *Maxfield v. Comm'r*, T.C. Summ. Op. 2006-27, No. 8075-04S, 2006 WL 354656, at \*3 (T.C. Feb. 16, 2006) (non-precedential) (holding that petitioners did not have reasonable cause for claiming improper deductions because their tax software depended on the entry of correct information and was not at fault for the error).

126. See PHILIP LEITH & AMANDA HOEY, *THE COMPUTERISED LAWYER* 315–17 (2d ed. 1998) (describing resistance to advisory and decision-support systems in the legal community).

127. See *id.*

128. See Darryl Mountain, *Could New Technologies Cause Great Law Firms to Fail?*, J. INFO. L. & TECH., Feb. 28, 2001, [http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2001\\_1/mountain](http://www2.warwick.ac.uk/fac/soc/law/elj/jilt/2001_1/mountain).

129. See Alan Cohen, *Data, Data Everywhere*, L. FIRM INC., Apr. 2007, at 16, 19–21, available at <http://www.law.com/jsp/PubArticle.jsp?id=1179392701422> (describing the efforts of law firms to profit from the electronic discovery process).

130. See John Buley, *Eight Things Keeping Law Firm Management Awake at Night*, L. PRACTICE TODAY, Nov. 2005, <http://www.abanet.org/lpm/lpt/articles/fin11051.html> (noting current price insensitivities, but predicting an increase in law firm price competition).

131. See Lauritsen, *supra* note 1, at 173–74.

132. See Mark Chandler, Gen. Counsel of Cisco Sys., Inc., Luncheon Address at the 34th Annual Securities Regulation Institute: State of Technology in the Law (Jan. 25, 2007), available at [http://blogs.cisco.com/news/2007/01/cisco\\_general\\_counsel\\_on\\_state.html](http://blogs.cisco.com/news/2007/01/cisco_general_counsel_on_state.html) (calling for firms to streamline legal processes by relying more heavily on IT and client self-help).

## VI. CONCLUSION

Although the legal profession already uses some computer technologies to automate law practice, and to store and retrieve documents electronically, there is a clear gap between the extent of adoption of sophisticated IT by other industries and by law firms. This gap has manifested itself as increasing costs and as unrealized efficiency gains. Fortunately, the research avenues surveyed by this Note, if realized, are likely to ameliorate these problems. New ways of constructing arguments, new methods for analyzing large sets of legal data, and new systems for representing that data will enable attorneys to reduce much of the repetitive waste they encounter in contemporary practice.

Although there are a number of potential barriers to the adoption of new computer technologies in law, it seems inevitable that the large profit margins commanded by law firms, and the comparatively repetitive nature of some of their work, will lead clients to demand that more of the legal process be automated, streamlined, and put in their control. In addition to making legal practice cheaper and more efficient, some of the tools described in this Note may one day aid in the dispersal of legal knowledge beyond the current bounds of the profession, to clients and the lay public. At the same time, new technologies promise to remove some of the drudgery from the practice of law, and to allow lawyers to focus on analysis of unsettled or ambiguous issues. Although it is difficult to predict which technologies will emerge, or when they will do so, one can remain confident that the adoption of such technologies will almost certainly bring significant changes to the practice of law.