



The Future of Legal Search: Meeting Lawyer Requirements by Delivering More Contextually-Sensitive and Relevant Results

Executive Summary

Research is one of the most vital and time-consuming activities in a lawyer's workload. While the legal information industry has innovated by improving search efficiency, the effectiveness of search largely relies on the research expertise of lawyers, paralegals and law librarians. The vast growth of available data adds additional challenges to the task of identifying the most critical and relevant information to a case or client matter.

With access to activities and data that a lawyer manages – including client information, current matters and ongoing client development opportunities – next-generation search algorithms should be able to understand lawyer research requirements and automatically parse legal information databases to extract relevant resources. The vast search history and usage patterns available from information service players can be leveraged to build intelligent search systems that can largely automate much of the tedious research work and present relevant results directly, improving speed, accuracy and relevance.

This white paper explores a scenario for constructing an automated search engine that parses complex legal information and returns more contextually-sensitive and relevant results that can be easily integrated with a law firm's existing knowledge management systems and workflows.

Need for Change

Legal research is the backbone for building a strong case. For decades, lawyers have been dependent on the duopoly of Lexis.com and WestLaw to provide vast information databases, combined with internal expertise and experience in searching, to uncover relevant results. While the industry has always focused on expanding coverage of legal information sources, it is ultimately the relative ease and accuracy of finding relevant resources that can generate maximum value for the legal community. Information service providers are providing options to allow lawyers to search more efficiently, but the overall results are still largely a measure of the lawyer's inputs.

The industry has operated in a "pay-per-search" or flat-rate subscription model, which has forced lawyers to be judicious in their use of search. While large law firms have often negotiated flat-rate subscriptions to large comprehensive databases, smaller firms and sole practitioners have felt the pinch. In many cases, the emphasis has been on paying for search rather than the results. Attorneys' clients pay for results, not the process. A similar practice can be adopted in the legal information search industry, with revenues derived from returned results – not the search process.

Market trends – such as the rise of alternative fee arrangements and initiatives like the Association

of Corporate Counsel's Value Challenge – have forced law firms to find efficiency while not sacrificing accuracy and comprehensiveness. As such, any ability that lawyers have to quickly obtain the

Social media, crowdsourced data and other sources of information continue to generate volume and increase complexity.

most accurate and complete set of results is likely to help them meet this efficiency goal and client needs.

Large law firms have built their own knowledge management databases and have huge volumes of work products available. Integrating these data sources along with the

public legal sources can maximize the collective knowledge for members of the firm.

Changing Industry Themes

While numerous developments have impacted information services players, several critical themes have emerged to truly transform the industry:

- The explosion of complex data.
- Changing consumption patterns.
- Conditions for business model innovation.

As the sheer volume of data increases, so does data complexity. Consumers of information now see the rapid rise of audio and video not only as part of multimedia sources but also as a primary means of delivering a variety of content types. As more and more courts digitize, some are going a step further. For example, the Illinois State

Supreme Court has eliminated print records of its cases, fully digitizing them but also providing near-real-time audio and visual feeds of court proceedings. Social media, crowdsourced data and other sources of information continue to generate volume and increase complexity.

Consumption patterns continue to evolve, even beyond the obvious impact that mobile devices have had on the industry. Users are increasingly comfortable with content formats other than text alone. User-

generated content, audio, video and podcasts are now considered relevant sources of information. Many users expect information services providers

to become a one-stop shop for not only all categories of legal content, but also any content that can have a bearing on a case.

Truly intelligent search has arrived. Examples include the iSeek capabilities currently in use at the University of Pennsylvania Medical School and Wharton Business School, as well as the semantic search capabilities of TotalPatent and other similar solutions.¹ Information services consumers expect integration into their workflow, such as Jigsaw Data Fusion's seamless connection with Salesforce.com and D&B 360's interoperability with Microsoft Dynamics. As consumption patterns shift, so does the expectation of information services providers to provide meaningful insights and contextually-rich content that is easily accessed, searched, found and applied.

With the rise of newer players in the market (i.e., Bloomberg Law) that are seeking to expand their reach through novel pricing models and growing price sensitivity among users, a push toward a pay-per-results model could emerge. Free content providers such as Google Scholar have also impacted pricing and revenue models. With new tools and smart content available, information services providers are able to offer micro-segmented and localized data on a scale previously not possible. All of these conditions have created opportunities and even mandates for innovative business models.

Collectively, these trends make clear the need for information services providers to innovate on their core search platforms and provide additional value-adds for the legal industry.

A New Look at Search

The current search model involves lawyers who use their intellectual capacity to determine appropriate keywords and specifying filters. Results are determined through strict keyword matching and use of synonyms. Leveraging search history, information search providers can start analyzing how lawyers actually search to build artificial intelligence tools for constructing queries based on cases on which a lawyer is currently working.

Factors such as the metadata for the case and the position of the lawyer, along with the stage of the proceedings, can be used to determine the lawyer's research objective. Search engines translate this objective into a search query, which can be run against the legal information database.

Leveraging search history, information search providers can start analyzing how lawyers actually search to build artificial intelligence tools for constructing queries based on cases on which a lawyer is currently working.

Search Model

Building the Search Framework

By analyzing search queries run by lawyers, as well as using standard dictionaries, it is possible to build a legal search ontology (see Figure 1). The following dictionaries can be utilized to do this:

- Legal terminology dictionary (a dictionary of legal terminologies that also shows similar and related legal terms).
- Practice-specific dictionaries.
- Industry-specific ontologies.
- Keyword dictionary based on keywords used in search.
- Corporate and legal directory showing major corporations, legal firms, etc.

Search Query Pattern

Search queries of successful searches can be analyzed to discover patterns for searches across various case types (see Figure 2). Patterns would cover the usage of filters like jurisdiction, date range, litigants, Boolean operators used between keywords and type of keywords used.

Deriving Context Sensitivity

To understand what a lawyer requires in his research, the system should be capable of identifying the context of any case. Deriving context

involves analyzing the pleadings to understand the legal issue. A case keyword list is generated by running the legal search ontology against the case documents to uncover keywords related to a specific case. The output will form the backbone of context. For example, for patent-related cases, the presence and absence of keywords typically associated with “patents” will be checked to form context (i.e., violation of patent, inappropriate licensing fees, objection to granting of patent, etc.). Product line and litigant involvement can be generalized to enhance context by using attributes such as industry and category of litigant (multi-national, government organization, small firm, individual, etc.). This will aid in finding resources relevant to this case category.

Deriving context involves analyzing the pleadings to understand the legal issue.

Metadata for each case will contain context and case information, such as case type, jurisdiction, industry, litigants, attorney, judges and stage of proceeding, along with keywords derived.

Search Process

Building the Search Query

The purpose of search entails:

- Identifying resources that are related to the context of the case.

Building a Legal Search Ontology

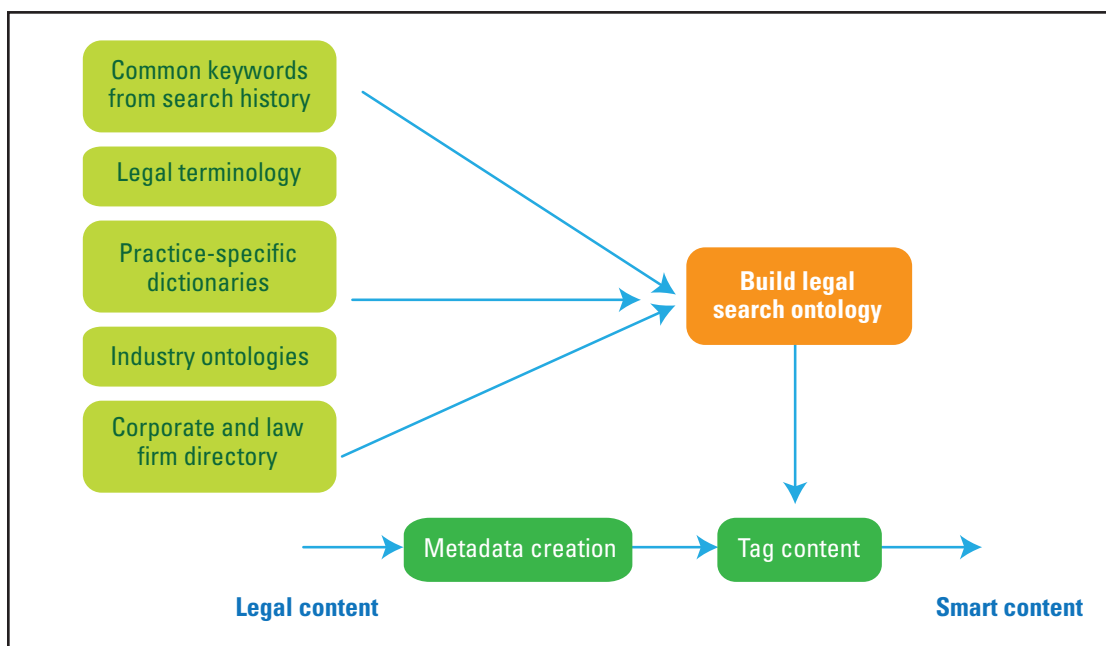


Figure 1

Using Pattern Recognition to Optimize Search Query Results

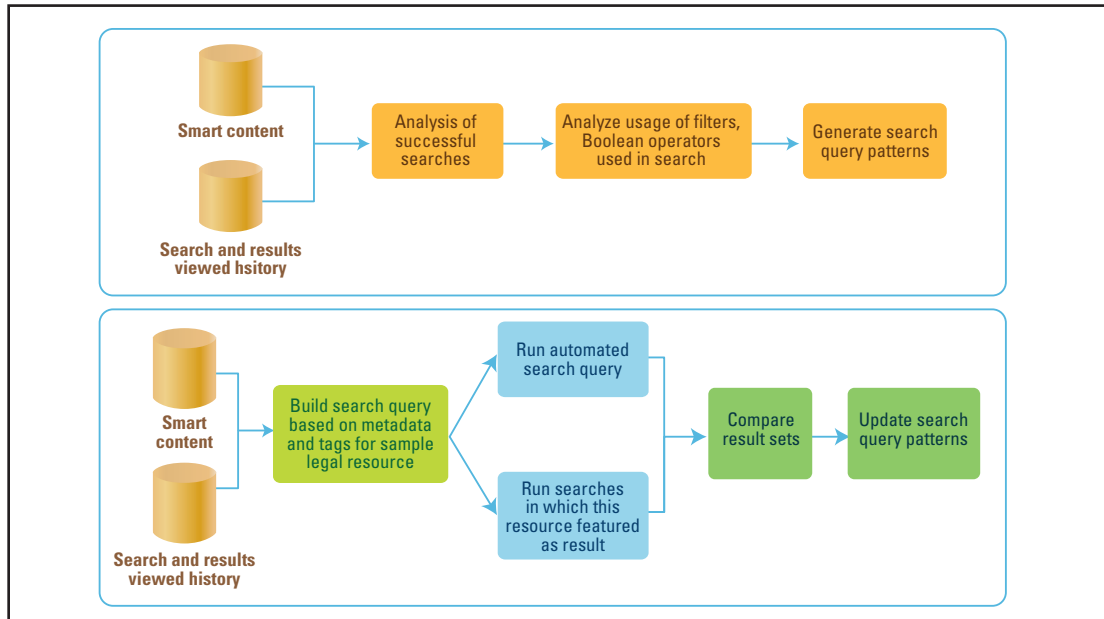


Figure 2

- Uncovering information related to litigants involved in the case.
- Discovering trends with potential ramifications.

When a lawyer indicates the case being worked on, the system generates multiple search queries to address various categories of search. Metadata from the case, including context, will form the building blocks of the search query. The search query will be formed by arranging these blocks in a suitable pattern, with a relevance score formula set based on search pattern history. This search query will be run against the database to determine results. The scoring formula will assign weights to various parameters based on how relevant they are to the search.

Resources matching the case context are considered most relevant. “Keyword matching” search results will be ranked at a lower priority. Keywords still need to be considered to ensure no relevant results are missed even though they may not match the exact context. A wider interpretation is required for trends. Search results will be classified into various categories during display to allow the lawyer to quickly drill down to what is required.

This system learns how best to search by continuously updating patterns based on user behavior (e.g., modification of the query by the user, deter-

mining which results presented in the search are selected, etc.).

Similar Results

By leveraging user history, it is possible to identify results that are similar in nature but do not meet the search criteria. When two or more results from a search are viewed, there is a high possibility that the results are related. By combing through the result view history, it is possible to identify groups of resources that reveal a high degree of similarity based on user perception. If one of the results matches the search criteria, there is a strong likelihood that other similar resources may be relevant. By presenting these resources as a, “You may be interested in these results” section, the lawyer is provided a broader range of potentially relevant search findings.

Fine-tuning for Optimal Results

In case of searches that generate very few or no results, the search query may need to be modified to generate results that are of lower relevance. Options include broadening date ranges, jurisdictions, industries involved and more lenient consideration of keywords.

Display of Results

Display results clustered around various views (filters) – source, industry, litigant, jurisdiction, date range, etc. – will lead to better organiza-

tion of results. Selecting a particular view will allow the lawyer to drill down into the results. Additional relevant resources that cover profiles, analytics, public records, etc. can be shown in a sidebar to deliver complete coverage. Integration of “similar results” and relevant work products from the firm’s knowledge management system, when possible, will deliver comprehensive results and ensure tighter integration of research with internal work products. Leveraging other online litigation products can display filings made by the firm, enabling additional productivity and effectiveness.

Color coding of results can further help lawyers select what is most important to them. Color-coding schemes can be used to indicate whether a result has already been viewed by the lawyer or another colleague, as well as previous involvement of the law firm or its client and whether the result has been recently modified.

Lawyers can also set alerts for any view displayed in the results page to allow for more timely and effective change tracking.

Modification of Search by the Lawyer

If a lawyer is not satisfied with the search run by the system, he can view the query run by the system and modify it. These modifications will be stored by the system and used to refine its search algorithm. Lawyers can also view options to set preferences for searches generated. They can specify which parameters to strictly ensure, which to ignore, as well as relative relevancies and threshold relevancies. They can also customize the views to match their priorities.

Extension of Search

Proactive Search

Lawyers are willing to pay a premium for timeliness of critical information and access to high-value expert content. Legal information services providers typically tend to provide

Color-coding schemes can be used to indicate whether a result has already been viewed by the lawyer or another colleague, as well as previous involvement of the law firm or its client and whether the result has been recently modified.

Automated Search Process

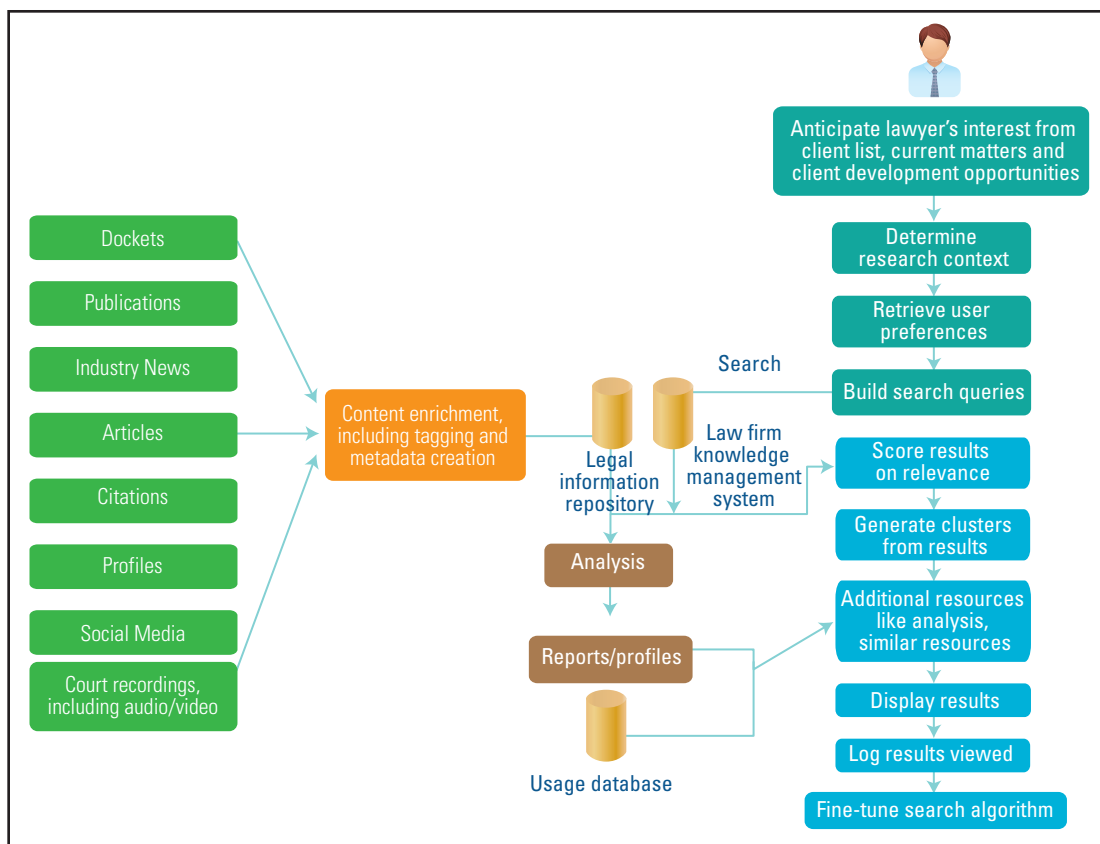


Figure 3

alerts on a daily basis, summarizing all the key changes that a lawyer wants to keep abreast of. Newer players, such as Law360, are addressing this gap by providing near-real-time alerts and expert opinion.² However, they still rely on lawyers explicitly configuring the system to follow alerts.

The intelligent search system can be used to proactively find information that is relevant to the lawyer. By knowing which cases a lawyer is

Proactive search is an ideal opportunity to highlight the value of paid content. By providing relevant free content and abstracts of paid content, the legal information industry can target upgrading of customers.

currently working on, as well as the lawyer's long-term clients, the system can search for new and interesting developments and deliver them in the form of real-time alerts.

A customizable home page that makes relevant content collections easily accessible provides a valuable service to lawyers by keeping them up to date with the latest happenings. Based on user-stated information and usage patterns, the page could emphasize different resources to solo practitioners, for example, than for a litigator at a large firm. The legal information industry divides content into "paid" and "free" categories. For long-term growth, information services providers need to increase the conversion rate of free content consumers to paid content consumers. Proactive search is an ideal opportunity to highlight the value of paid content. By providing relevant free content and abstracts of paid content, the legal information industry can target upgrading of customers.

Educational publishers give students the option to buy portions of content that are tailored to their needs. Licensing rights also include access to content that they own in perpetuity or for a specific duration of time (i.e., a semester). Similarly, rights to content assets can be made available through bundles or collections that are convenient to the end user. Proactive search can enable creation of such bundles that will be available for specific periods of time (e.g., duration of a trial).

Proactive search is also an ideal way to consume research when using smartphones and tablets. This reduces the need for users to physically type keywords and search; instead, they only need to navigate through returned research content.

Litigation Workflow Platform

The search platform can be extended beyond the traditional research space to cover other areas of the legal value chain. Integration with online filing systems can help lawyers quickly retrieve work products from their firm with a strong bearing to the motion currently being filed. The platform can also be extended to help people who are searching for a lawyer. Based on the keywords and other details entered, the system can retrieve a list of lawyers who have worked on similar types of issues by searching for cases with similar context and extracting lawyer names from the metadata. It can also show resources such as lawyer profiles, which can help users make smarter decisions on hiring lawyers and verdicts for similar cases to analyze the possibility of winning cases.

New Social Sources of Content

The proliferation of user-generated content, videos and social media is challenging the current legal research system. Most of these new sources of content can help lawyers gauge public perceptions that can help them win over juries, such as video proceedings of court activities that can help them analyze nuances such as non-verbal communications. Given that Twitter, Facebook and other social media postings may become vital evidence in many categories of cases (e.g. discrimination actions), lawyers need to be aware of such postings made by their clients or opposing parties.

The current challenge is how to help legal professionals easily discover such information. Much of this content is fragmented across the Internet with very little value-add in terms of making it discoverable and easier to analyze. Feeding these sources of information into the context-based search framework will enable the co-mingling of new, relevant multimedia content alongside traditional sources of legal information, providing a one-stop research solution.

Collaborative Search

Search results and alerts can be used to populate research repositories for various contexts. These repositories can be shared with colleagues, allowing them to view results, suggest new sources for the repository and add comments. Repositories allow for the tracking of who has viewed and downloaded results and also partitioning results into subsets that can be reviewed by different members of the team. A colleague can also follow a repository via alerts that are sent to subscribers

when new results are added. Repositories can be stored in the firm's own knowledge management systems or service provider's cloud infrastructure, allowing easy access from any device.

Lawyers can share abstracts of user-generated content and other results from the repository with acquaintances in their professional network to get advice on source reliability. Unreliable results can be discarded from the repository. User-generated ratings can be used to guide the relevance of results.

Future of Work

For a long time, legal research products have been targeted for legal professionals, government institutions and law schools, and the products have been developed in silos, away from other tools used by the legal profession. Legal search needs to be better integrated with other aspects of lawyers' and other legal professionals' daily tasks and also address the needs of non-professional users.

The Non-Legal Community

The "do-it-yourself" trend is invading the legal space. Rather than approaching a lawyer at the first instance of trouble, many people now conduct initial research themselves before contacting a lawyer. To cut costs, many individuals often draft their own contracts, wills, etc. rather than using lawyers.

Many of these individuals rely on free sources of legal information such as Google Scholar and Fastcase. These players offer little value-add but are often seen as the only affordable sources of information. Incumbent legal information service players such as LexisNexis and WestLaw now offer free versions of their search products with limited functionality and content. However, they have had limited success in converting free users to their paid products, which is their ultimate aim. Better value propositions such as pay-per-result and assistance in discovery of relevant results can improve conversion rates. A new intelligent search framework can help them in this regard. Further integration of search with other products – such as legal form downloads, blogs for expert legal opinion, community forums for discussing legal topics and lawyer leads – can provide a one-stop solution for addressing the needs of this community.

Legal Community

Device Agnosticism

The shift in research from print to electronic has drastically altered the way legal research is being

performed. A second shift is now on the horizon in which the medium of consuming research will expand from PCs to encompass alternate electronic devices, such as smartphones, tablets and e-book readers. The legal search industry needs to address content distribution and accessibility limitations of their products to ensure an optimal experience across electronic devices.

Integration Across Litigation Products

Large legal information services players like LexisNexis and WestLaw also develop other products targeted at lawyers (e.g. case management and analysis, CRM and practice management tools). However, development is often conducted in silos and provides limited opportunities for collaboration among lawyers.

Ideally, a single-sign-in, cloud-based solution that provides access to various tools and ensures maximum integration of research and case data with litigation tools will benefit lawyers the most and also help to attract users and keep them loyal to one platform.

Current examples of collaboration services – such as Lexis for Microsoft Office, which combines Outlook, Word and SharePoint – that provide collaboration tools and include search functionality for desktop and internal document management systems need to be extended to cover the entire litigation platform.³

Personalization

Legal search providers need to personalize content based on stated interests (registration), electronically recorded habits and aggregated data. News needs to be available in real-time for lawyers to be aware of the latest happenings that can affect their clients. This includes fully customizable U.S., local and other news, along with the latest business and professional news covering practice-specific articles, client news, verdicts and noteworthy cases.

Original content

Search providers need to make available original content, apart from acting as information aggre-

Better value propositions such as pay-per-result and assistance in discovery of relevant results can improve conversion rates.

Ideally, a single-sign-in, cloud-based solution that provides access to various tools and ensures maximum integration of research and case data with litigation tools will benefit lawyers the most and also help to attract users and keep them loyal to one platform.

gators, to maintain their relevance to lawyers, as well as tackle the growing threat of free legal information aggregators. The combination of

in-house journalism, licensed content feeds and user-generated content can ensure that better information is delivered to users. Certified experts can provide no- or low-cost blog and video blog content. Search providers can also build online communities of legal professionals who can link to each other and also share discussions, best practice tips and network.

Ultimately, search providers must engender loyalty by offering more utility in

one place, deriving value through social media resources and creating a captive base to cross-sell other products and services.

Free Users

Legal search providers need to convert free users into a loyal and captive audience that becomes familiar with their brand and frequently returns through e-mail alerts, newsletters and social media updates. This well-indexed and constantly refreshed content creates a larger Web presence

and search engine optimization advantage to further promote usage and minimize the reliance on advertising spend to build brand. Revenue can be derived through advertising, pay-per-results and conversion to paid services. Frequent trial offers can be used as a promotional tool to provide partial functionality.

Similar user experience for paid and non-paid users with more functionality and additional access to content

for paid users needs to be provided. For example, a certain number of placeholder content boxes might always be persistent on the screen, even though the services have not been activated. If users select a content module to which they're not subscribed, there will be a call to action on how to subscribe and activate it. Similarly, this may also

be an effective way to acquire customers for limited-use sites such as EZ Law that are generally used only as needs arise.

Addressing Lawyer Concerns

There are always misapprehensions regarding new technology whenever it is introduced. There is a high probability that lawyers and librarians may be skeptical of the system's ability to judge what is relevant. To build initial accuracy of system while not compromising the quality of results provided to lawyers, such systems should debut first in law colleges, where law students can rate the results provided and suggest modifications to the search engine. These recommendations can be used to tune the search system before it is rolled out to a wider audience. This approach may lead to quicker acceptance by the legal fraternity.

Another issue that can hinder adoption is the legal staff's perception that their years of expertise in research are being rendered obsolete, which might make them resist adoption of such platforms. Librarians and other legal staff may also feel threatened by automation in their traditional research domain. To address these concerns, search automation can be positioned as a tool that can help legal staff reduce time spent in mundane search activities and focus more on understanding and analyzing results to generate maximum value.

Final Arguments

With the growing advancement of search technology, it is possible to automate large parts of legal research. Leaving the tedious work of searching to computer systems allows lawyers to focus on building a strong case. The explosion of complex data makes this capability not only more relevant but also more critical to ensure that lawyers arrive at the right answers quickly and efficiently.

Leveraging search and usage history can result in the construction of more effective search systems that anticipate what a lawyer requires and quietly work in the background, delivering relevant information directly.

In the future, the entire process of search may, in fact, disappear from a lawyer's workload. Information systems will continuously keep track of changes and download relevant information directly to the lawyer's work folders.

Ultimately, search providers must engender loyalty by offering more utility in one place, deriving value through social media resources and creating a captive base to cross-sell other products and services.

Search automation can be positioned as a tool that can help legal staff reduce time spent in mundane search activities and focus more on understanding and analyzing results to generate maximum value.

About the Author

Ambika Sagar is an Associate Consultant within Cognizant's Information, Media and Entertainment Consulting Practice. She has previously worked at online job search portal Naukri.com, where she focused on improving the search experience for recruiters. Ambika is an information technology engineer with an MBA in marketing and systems from XLRI School of Business and Human Resources. She can be reached at Ambika.Sagar@cognizant.com.

About Cognizant

Cognizant (NASDAQ: CTSH) is a leading provider of information technology, consulting, and business process outsourcing services, dedicated to helping the world's leading companies build stronger businesses. Headquartered in Teaneck, New Jersey (U.S.), Cognizant combines a passion for client satisfaction, technology innovation, deep industry and business process expertise, and a global, collaborative workforce that embodies the future of work. With over 50 delivery centers worldwide and approximately 118,000 employees as of June 30, 2011, Cognizant is a member of the NASDAQ-100, the S&P 500, the Forbes Global 2000, and the Fortune 500 and is ranked among the top performing and fastest growing companies in the world. Visit us online at www.cognizant.com or follow us on Twitter: Cognizant.



World Headquarters

500 Frank W. Burr Blvd.
Teaneck, NJ 07666 USA
Phone: +1 201 801 0233
Fax: +1 201 801 0243
Toll Free: +1 888 937 3277
Email: inquiry@cognizant.com

UK Headquarters

1 Kingdom Street
Paddington Central
London W2 6BD
Ph: +44 207 297 7600
Fax: +44 207 121 0102
Email: infouk@cognizant.com

India Operations Headquarters

#5/535, Old Mahabalipuram Road
Okkiyam Pettai, Thoraipakkam
Chennai, 600 096 India
Phone: +91 (0) 44 4209 6000
Fax: +91 (0) 44 4209 6060
Email: inquiryindia@cognizant.com
