TIG Final Evaluation Report

Grantee name: Legal Services of N. California   TIG grant number: 07505
Submission date: July 28, 2009 (revised)   Approval date: July 29, 2009

Contact person: Brian Lawlor   Phone: (916) 551-2119
Email address: blawlor@lsnc.net

I. Project Goals and Objectives

The overall project goal was to improve the effectiveness and efficiency of client services by enabling Legal Services of Northern California (LSNC) staff to share and locate documents throughout the organization via a secure, enterprise-level, multi-office "knowledge content" system based on Google Search Appliance (GSA) and Microsoft SharePoint technologies. LSNC broadly characterized this technology initiative as "The Findability Project." The 18-month project term was from January 2008 through June 2009.

The project had four core objectives:

Objective 1: Modify the LSNC network infrastructure to incorporate the hardware and software required for implementation of an effective and efficient program-wide search capability.

Objective 2: Develop document repositories using methodologies and protocols that enable users to easily add and retrieve shared content.

Objective 3: Implement a knowledge content system relying on the Google Search Appliance to enable staff throughout the program to easily and efficiently locate and retrieve shared information.

Objective 4: Foster replication of the system by publicizing, via a "Findability Project" development website and other media, the system's technical components, explanations of how the system was implemented, and its impact on LSNC's operational effectiveness and other related "lessons learned."

No significant changes were made to the overall project goal and the stated objectives.

II. Evaluation Data and Methodologies

Objective 1: Network infrastructure changes

The changes to the LSNC network infrastructure required for implementation of this project were based on the basic hardware and software technical requirements of the Google Search Appliance (GSA) and Microsoft SharePoint. At the beta phase of this

project, a GSA consultant conducted an onsite evaluation of those components and, at the system level, conducted a series of GSA network-communication and search-query test challenges to confirm intended functionality of the new system.

Objective 2: Document repositories, policies and protocols

This objective involved making decisions about how the shared document repositories should be organized, and once organized how easily staff could navigate to the document locations to add, remove or retrieve files. To evaluate these options we went through an extensive vetting process among senior advocates, managing attorneys, office managers, and ultimately all staff in all affected offices. Surveys were also conducted after trainings for the two test-bed offices to determine whether staff understood the organization of the new system.

Objective 3: Enterprise search implementation

Test-bed installations of the GSA and all other hardware and software requirements were completed in the Sacramento and Chico offices in January 14, 2008 and April 4, 2008, respectively. Installation and comprehensive integration of the remaining six LSNC core offices into the new system was completed on August 13, 2008. To evaluate this new enterprise-level system, we did four things:

First, for the two test-bed locations, in October 2008 we conducted an onsite training for the Sacramento office and a remote training via LegalMeetings for the Chico office, after which conducted a survey to evaluate whether staff:

- understood how to access the new system.
- understood how to search for, locate and retrieve documents.
- succeeded or failed in their effort to find what they were looking for.
- found the new system easy or hard to use.

Of 31 persons in those two offices, we received 17 survey responses.

Second, at near completion of the project, we conducted onsite training of the new system in the Redding office on March 25, 2009, after which we conducted a survey comparable to the one used earlier with the test-bed locations. Of the 12 attendees at that meeting, six survey responses were submitted.

Third, after making modest additional technical changes to the system, we conducted three program-wide remote trainings via LegalMeetings on May 26, 27 and 28, 2009. These trainings were promoted and scheduled to encourage attendance by all interested staff program-wide. After completion of those trainings, we conducted a program-wide survey to cull additional user response, similar to what had been sought in earlier surveys, but requesting additional feedback about:

- user ratings of the "file type" and "special collection" vertical filtering tools added to the interface (permitting users to narrow their search results to a slice or subset of the original results).
- user ratings of the GSA OneBox specialized database search feature for locating individual and office personnel listings.
- user ratings of the impact of the new system on the overall effectiveness, efficiency and quality of one's work.

Of 120 employees program-wide, we received 35 survey responses.

Fourth, upon completion of the project, it was reviewed remotely on June 22, 2009 for a technical and best-practices assessment, and follow-up recommendations, by a four-member team of GSA specialists with the Google Enterprise deployment team.

Objective 4: The Findability Project development site

The Findability Project development website was launched and promoted publicly on July 21, 2008. To evaluate the effectiveness of the site over the remaining 12 months of the project life, we compiled data using the site's native WordPress publication metrics as well as Google Analytics and the Google FeedBurner service, to track the following:

- tutorial and other special content pages published.
- additional informational posts published.
- approved comments to various published posts.
- how many people visited the site.
- the number of pageviews of site content.
- the "bounce rate," i.e., how "sticky" the site was such that visitors proceed to view additional pages after landing at the site.
- how frequently people viewed the site via an RSS feed.

III. Summary of Major Accomplishments, Recommendations and Future Steps

The overall project goal and all stated objectives were completed successfully. With completion of this project, LSNC now has in place a well organized, practical, cost-effective, user friendly knowledge-content system that LSNC staff overwhelmingly report as helping them get their work done on behalf of clients more effectively, efficiently and with improvement in overall quality.

Major Accomplishments

Among the most significant accomplishments of this project are the following:

- LSNC completely rebuilt the network and software infrastructure in its eight core offices (the flagship Sacramento office, as well as its remote office locations in Auburn, Chico, Eureka, Redding, Ukiah, Vallejo and Woodland) to support this

project. Component implementation program-wide included varied installations and configurations of Windows Server 2003, SharePoint Server, Microsoft ISA Server 2006, Google SharePoint Connector and, of course, an enterprise-level Google Search Appliance (GSA).

- To improve overall findability of documents within the system, LSNC created a simple, easily understood 29-category structural taxonomy for organizing the organization's content, based on the 10 basic LSC legal problem codes, complemented by 19 other categorizations, covering every major aspect of the organization's substantive and administrative work. These same 29 taxonomic categories have also been adopted as optional metadata "tags" that can be added to files in a way that enables users to search for a document across one or more categories.

- Formal policies and protocols, as well as recommended "best practices," were created, thoroughly vetted and adopted to standardize throughout the organization: how files are named; where they should be located in the structural taxonomy; where and how individual user directories should be named and located; what methods should be used to add or remove files from the system; and how to process and convert hard-copy documents to digital format for addition to the system.

- Once this new system was fully deployed, web accessible and usable by all LSNC staff, all LSNC's designated knowledge content was easily "findable" throughout the organization, i.e., users in all offices could rely on enterprise-level Google search functions and filters to locate and retrieve the information they need to do their work. Among the major highlights of this new system are the following:

  o A customized, universally web-accessible, secure "search" user interface (UI) was built as a single point-of-search entry for locating, viewing and/or downloading shared documents throughout the organization.

  o Any search results generated by the system can be vertically filtered to narrow the results to include only those of a particular file type (e.g., PDF files) or only part of specific collections (e.g., administrative forms, pleading exemplars, etc.).

  o Google "OneBox" modules — special GSA-driven database queries that return the data searched for as a special search result at the very top of the search result page — have been integrated into the search UI, enabling users to search instantaneously for individual or office-specific personnel contact information; a similar OneBox module has been integrated to search for and return special results for all administrative and case forms used throughout the organization.

- All authenticated users throughout LSNC can now search and/or navigate to all content that is targeted by this system. The system integrates all content in all designated LSNC private and public collections, including:

  o all official LSNC intranet content (policies, protocols, forms, administrative and case manuals, personnel and other human resources information; development and fundraising resources; technology-related information, etc.).

  o a peer-reviewed collection of recommended exemplars of pleadings, memoranda and other types of commonly used advocacy documents.

  o a comprehensive collection of supporting documents for all impact cases and projects.

  o all content at all LSNC public web locations, including its LSNC.net flagship site, the Race Equity Project, the California Food Stamp Guide, and all content at all of LSNC's five other special program websites.

  o hundreds of thousands of other user-generated, high-value character/text-based documents (WordPerfect, Word, Excel, PDF, PowerPoint, RTF and TXT files) to be organized taxonomically in shared document repositories.

- Multiple surveys of LSNC users confirmed that the new system is overwhelmingly well understood, well organized, and enables them to locate organizational content efficiently, effectively, and with significant gains in the overall quality of their work. Among the major conclusions of the surveys are the following:

  o 92% said the organization (structural taxonomy) of the shared document repositories made sense.

  o Between the initial phases of the project and its completion, the overall success rate of users in locating the information they were looking for improved from 73% to 90%.

  o Based on a four-star rating system, 89% of respondents rated the system at either 3 or 4 stars for its enabling them to more easily, quickly and efficiently find information within LSNC; 82% rated the system at either 3 stars 4 stars for its enabling them to be more effective at doing their work; and 71% rated the system at either 3 or 4 stars for improving the overall quality of their work.

- The project included creation of a public web development site, called The Findability Project, to offer transparency to the larger legal services community about the project, while facilitating understanding and replication of its goal and objectives. Over the last 12 months of the project, it regularly published commentary, tutorials and analysis of the planning, technical details, and search

concepts at the heart of the project. During that period, the site had over 4,400 visitors sessions and over 20,700 pages views. The site also averaged 600+ feed pulls per month.

- The Findability Project drew the attention of the [Google Enterprise](#) deployment team, which has requested to profile the project as an innovative, successful non-profit use of Google technologies.

Major Lessons and Recommendations

- Process and transparency are vital to the success of a project of this scale and depth of impact on an organization. The process of vetting major project elements, design and implementation, as well as vetting proposed practices, protocols and best practices, were essential to making a project like this work for all involved.

- It is critical to a project of this scale and complexity to retain a professional consultant who has real-world expertise in information architecture and related systems, expertise in implementing your chosen enterprise search platform, and can realistically assesses the organization's needs and resources for successful completion of the project.

- The Google Search Appliance offers superior advantages over other enterprise search options because, by design, the GSA works seamlessly with Google Docs, Google Sites and Google Analytics, web-based applications commonly used in legal services and other non-profit organizations.

- An organization should anticipate there may be particular technical challenges with integration of its case management or other custom content system into the enterprise search platform.

- Designing and implementing a comprehensive, fully searchable "knowledge content" system is anything but a "turn key" operation. However, there are a range of ways legal services program can begin to build in-house experience and expertise with Google search technologies and the local office and organization-wide, such as Google Desktop Enterprise, Google Apps (especially Google Sites and Google Docs); and more capable and customizable options like the Google Mini and the Google Search Appliance.

IV. In-Depth Analysis of Accomplishments

LSNC has successfully designed and implemented a well organized, practical, cost-effective and user friendly knowledge-content system that the overwhelming majority of staff report as helping them get their work done with gains in overall effectiveness, efficiency and quality. Among the most significant accomplishments of this project, corresponding to particular objectives, are the following:

Objective 1: Modify the LSNC network infrastructure to incorporate the hardware and software required for implementation of an effective and efficient program-wide search capability

Basic Hardware and Software Specifications

LSNC completely rebuilt the network and software infrastructure in its eight core offices (the flagship Sacramento office, as well as its remote office locations in Auburn, Chico, Eureka, Redding, Ukiah, Vallejo and Woodland) to support this project. This aspect of the project was fairly straightforward, simply involving component implementation of the basic hardware and software required by the technical specifications of the project, including:

- **Windows Server 2003 —** this is the software platform backbone, so to speak, of the new system. All local LSNC office-specific network shared file servers were built out using Windows 2003 Server, which allows installation of SharePoint Server 2007, the open source Google SharePoint Connector and Windows Server 2003 Active Directory, described below. Windows 2003 Server is a robust secure server which allows local domain and subdomain user authentication for multiple locations. It also provides centralized authentication for SharePoint access as part of the universal, shared domain login available to all LSNC offices.

- **Windows Server 2003 —** certified ASUS PM52-M motherboard — if you're going to build a platform suited to working with Windows Server 2003, it is not just a matter of purchasing a server that nominally meets basic system requirements. You also need to ensure it is certified to do just that. It is not the only such option, but the ASUS PM52-M motherboard is so certified. There are numerous other Intel chipsets that are not.

- **SharePoint Server 2007 —** SharePoint Server 2007 has been installed on a single server in the flagship LSNC Sacramento office. It has an essential role in the new, overall network infrastructure as the core, central server that communicates throughout the enterprise with all the local office servers, and provides them with the networked locus to the shared taxonomic structure, which resides on the SharePoint server. While it is not the only such server option that works with the GSA, SharePoint affords distinctive advantages because it includes document management features for sharing and managing documents and other types of files across the enterprise, as well as a practical method for adding metadata to shared content.

- **Microsoft ISA Server 2006 —** the Microsoft Internet Security and Acceleration (ISA) Server allows offsite access to content returned in search results by providing an Active Directory authenticated external web interface.

- **Google SharePoint Connector —** available as an open source project at the Google Code site, the Google SharePoint Connector provides a seamless

connection between the GSA and the Windows and SharePoint servers. This connection uses Active Directory authentication for managing content permissions as well as the interface for the GSA to access and crawl all the domain content.

- **Google Search Appliance (GSA)** — the Google Search Appliance is, quite literally the brain, the cerebral cortex, the heart and the soul of the Findability Project. It is a high-end computer with the capacity to provide enterprise-level search capabilities for crawling and indexing any targeted content within our organization as well as external sites. If the organization has permission to a site, the GSA can crawl and index it, and then return search results in different, highly customizable ways.

A related technical issue with the project's rebuilding of these network services was the decision to retain Microsoft Office 2003 as a default configuration on all LSNC desktops, rather than upgrade to Microsoft Office 2007. At the time this project was implemented, we discovered that a key feature in Microsoft SharePoint — the option enabling users to add metadata to a file — worked reliably and predictably with the Microsoft Office 2003 suite, but not with the 2007 version. With this metadata feature activated via SharePoint, users are able to use the "Save" dialog in all MS Office desktop applications (Word, PowerPoint, etc.) to directly add metadata to any file created or edited within the application and save it to any shared document repository location. (The Appendix includes illustrative screenshots of how the metadata option works)

Beta testing and evaluation

Building out the new infrastructure required by this project was well within the technical competency of the LSNC IT staff. However, the project still required special expertise from a GSA consultant to provide both remote and onsite assistance with optimizing the system configuration and thoroughly testing it. LSNC retained Michael Cizmar and Associates (MC+A), a prominent Chicago-based information systems consultant recommended by Google Enterprise.

LSNC first built out and configured its Sacramento office as the initial alpha test location in January 2008. We then spent the next few months experimenting with the GSA, exploring its various features, validating the effectiveness of its crawling and indexing protocols, and testing various configuration options. Because of the immediate success of the Sacramento installation, we proceeded to install and configure the system for use by the Chico office in April 2008.

It was only after our doing so that we arranged for the GSA consultant to come onsite to test and evaluate the new system in the Sacramento office location, and conduct remote location testing of the Chico installation. At this juncture we considered the project to be at a beta phase. The consultant came to the Sacramento office for two days, on May 22-23, 2008. He conducted a systematic review of all the hardware and software elements involved and then conducted a series of rigorous GSA network-communication and search-query test challenges to confirm intended functionality of the new system for

both test-bed locations. After the beta test-bed installations were thoroughly tested over this two-day period, there were only minor tweaks suggested and made to the system by the consultant. We encountered no significant problems with the system set up in the Sacramento or Chico office locations.

At this juncture, the basic system was stable and working as intended, although not all intended GSA content targets were in place. It took several months to complete the rollout of the hardware and software changes to the remaining six LSNC core offices, but that process was completed on August 13, 2008 with the final implementation in the Redding office.

Objective 2: Develop document repositories using methodologies and protocols that enable users to easily add and retrieve shared content

An enormous amount of collective effort went into the process of sorting out how best to "organize" LSNC's "knowledge content" as part of this project. The project posted a series of very substantial articles at its public development site detailing how LSNC sorted out the theory of taxonomy for such a system, how to make sense of such theory against the practical realities of a legal services program, and what conclusions LSNC reached about how best to proceed in organizing its content. (The Appendix includes three taxonomy articles published at the project development site.)

Creating the structural taxonomy

To assure the organization of LSNC's knowledge content makes inherent, practical sense for its users, and works effectively with the Google Search Appliance, LSNC distilled its approach to structural taxonomy down to three elements:

1. The directory structures need to be a hierarchical or "top-down" organization of simplified, familiar categories.

In the broadest sense of "organizing" things on a file server, and how that same "organization" is reflected in page menus or page navigation or dialog boxes, users need to know where they are and what the folders or subfolders mean. Lawyers, by training and practice, work in an especially pronounced hierarchical environment. (Can you say, "I, II-A, …") The practical truth is that almost everyone in a law office environment organizes their work in some hierarchical fashion, and this is the most common way in which most people organize things, lawyers and non-lawyers alike.

2. Names for content folders, subfolders or categories need to be consistent with the shared vocabulary of your organization.

When working out the naming conventions for folders and subfolders, it was important to focus on commonly understood, familiar shared vocabulary or terminology. From the perspective of the GSA, the particular names, as such, of directory folders or subfolders are of no consequence. The GSA does not care what you call things. The GSA will ferret out the content wherever it resides, regardless. However, the keywords used when doing

a search via the GSA do matter, and the GSA will very effectively search for, and give significant weight, to keywords that appear as part of directory structures. So it matters how one names the directories in a structural taxonomy.

LSNC adopted the most conventional names for its directories it could come up with, including the LSC substantive problem code categories, which comprise roughly one third of the directories on our shared document repository. The LSC organization of legal topics is what field programs, after decades of use, know and understand. The LSC categorizations are functionally part of the "shared vocabulary" of our organization, and its use offers consistency with how other organizational information and data is handled, most notably client case data.

3. "Lean toward a broad-and-shallow rather than narrow-and-deep hierarchy."

That's a quote from Peter Morville's classic work, [Information Architecture for the World Wide Web](). That observation is consistent with the advice our GSA consultant gave us, namely, do not to go more than two levels down below the top-level directory, and if you can, go only one level down.

The rationale was two-fold: First, the more subfolders you have, the less likely users will locate or use content in those folders whenever they are navigating the directory structure, in whatever form it is viewed. From the user side, a deeper vertical hierarchy actually reduces findability.

Second, from the GSA side, deeper hierarchy does little or nothing to improve search results. While the search algorithms baked into the GSA exploit the URL path at the directory and subdirectory levels to improve search results, having a third or fourth or more levels does essentially nothing to improve those results. There's no harm to doing so. It just doesn't help the search significantly.

A counterpart to this issue is the importance of striking a balance. By going broad-and-shallow, one gets the practical advantage of being able to add content without the need for major restructuring. Assuming one has figured out a set of top-level directories that pretty much covers, in a broad sense, the content users will want and need to search for, from there on out the organization can focus on adding content below that level, as warranted. But if one goes too broad, from the user side, things get more cumbersome and impractical. Whether the users are advocates or office managers or volunteers, it is more practical and useful if they can visually and cognitively understand the organizational scheme, i.e., the structural taxonomy. So it needs to be broad enough to cover the bases, but not so broad that it becomes incomprehensible.

Based on these considerations, LSNC modified its originally proposed 42 top-level directory structure with as many as three levels down, and simplified it to its present 29 top-level directory structure, each only going no more than one-level deeper. (See the Appendix for a complete listing of the structural taxonomy adopted.)

The institutional process for identifying and resolving how the shared document

repositories should be organized — and once organized how easily staff could navigate to the document locations to add, remove or retrieve files — involved multiple steps: Designated senior advocates and office managers (1) proposed a structural taxonomy for organizing documents; (2) proposed additional private intranet and public web content as targets of the search appliance; and (3) proposed protocols and "best-practices" for the addition, removal and naming of documents in the shared repositories. The proposals were then (4) circulated to all office mangers for review and comment; and then (5) revised further upon consensus at a program-wide meeting of office managers. Finally, (6) all of these proposals were vetted with all managers in all offices, and (7) eventually with all other staff, and then revised and finalized in response to that feedback.

In addition, after implementation of the Sacramento test- bed installation, (8) we then sat down, one-on-one, with every advocate in that office to demonstrate and explain the structures and how they relate to their individual work files and the larger enterprise search project. We then conducted office-specific trainings for both the Sacramento and Chico office test-bed locations, after which (9) we conducted surveys to evaluate, among other things, whether staff in those offices knew how to locate or navigate to the shared local-office and program-wide document directories, and whether they understood how the shared document repositories and other intranet content were organized.

LSNC received 17 total survey responses from those two office locations, and 12 responses to specific questions about how the content was organized. One of the 12 said they understood "some but not all" of the directory names. Of the remaining respondents, 92% reported that they understood the organization scheme adopted, either reporting that they "get it" (50%); the organization "makes sense but I would do some of them differently" (17%); or the directory structure "makes perfect sense" (25%).

Because of this expansive, inclusive process, we determined we could significantly reduce or consolidate the horizontal design of the taxonomy from 42 to 29 categories, and more significantly reduce the vertical design from three levels down to simply one.

There was one significant category addition made: Creation of a dedicated directory for "Trainings" broken down by provider name, a change that came from advocate feedback.

Adopting file-naming conventions

How files are named within a shared system like this directly impacts their "findability." For perspective, if one were randomly to look at the files in advocate staff directories, one would likely see individualized (albeit typical) patterns in the file names: Advocates use both directory structures and name files in a way that makes them "findable" for them later, if not for others. The challenge with this project was to adopt file naming conventions that make sense for everyone, not just an individual user, while at the same time enhancing the findability of the files as part of enterprise search.

After studying this issue, the file naming convention we adopted involves varied use of the following elements:

[draft/final] [document type] [party/case] [subject] [date] [file extension]

For example, a file may be named "final_1085writ_valles_calworks.doc".

At the system level, we learned there were several practical factors impacting file naming conventions:

- Using spaces between words in file names creates file transfer problems when moving the file from one server to the other.

- Using spaces creates readability problems when viewing the path of the file in a GSA search result, because the GSA normalizes the URL by inserting special characters wherever the file has a space in its name. For example, "draft writ valles" is normalized by the GSA to the harder-to-read "draft%20writ%20valles".

- File names with underscores are easier to read than files with dashes. Although only based on anecdotal testing, our experience with users is that they report it to be easier to read file names with underscores between the words, as opposed to file names with dashes. For example, "draft_writ_valles" is easier to read than "draft-writ-valles".

- If one uses underscores, a linked file name in a search result is easier to read because, as a link, the file name appears underlined, so words appear as if they have spaces. For example, "<u>draft_writ_valles</u>" is easier to read than "<u>draft-writ-valles</u>".

(The Appendix includes the final version of the "file-naming best practices" adopted by LSNC as part of this project.)

Adopting document policies and protocols

Based on our experience working on this project, we thrashed out the practices and protocols detailed in the two memos attached to the Appendix describing "Best practices and protocols for GSA documents" and "Organization of advocate-user directories." Hopefully these memo are self-explanatory and offer practical logic and straightforward approaches for individual users or designated support staff to add files to the shared document repositories.

Among the most practical observations in these memos, however, is addressing the common but incorrect perception that one needs to save a document to "the" correct directory, as opposed to "a" correct directory, however it is done. For example, in California, a Section 1085 writ involving a legal challenge to a state welfare policy, logically, could be added to a directory with civil practice documents or to a directory on income maintenance, either of which is appropriate and in either case will be located by

the GSA.

The memos also attempt to address some of the practical realities and limits of a non-profit, legal services work environment. LSNC has neither the resources nor motivation to micro-manage how users organize their own file directories. Life is too short. But as detailed in the "advocate-user directories" memo, we do now require all LSNC staff to have a user-specific, user-named directory, and that the name used be the user's full name. The primary motivation for this requirement is a practical need to standardize directory-name conventions throughout the organization, so that the location and targeting of files is predictable, manageable and findable, as needed, by the GSA. It also provides obvious, practical improvement in findability whenever a user is using a file manager program or an application's file dialog to navigate to a particular user's directory, since the name of the user is predictable since the name is standardized.

Objective 3: Implementation of organization-wide search capability enabling all staff to easily and efficiently locate and retrieve shared information

Search Features

Once this new system was fully deployed, web accessible and usable by all LSNC staff, all LSNC's designated knowledge content was easily "findable" throughout the organization, i.e., users in all offices could rely on enterprise-level Google search functions and filters to locate and retrieve the information they need to do their work. Among the major highlights of this new system are the following:

- Using features available within the Google Search Appliance, a customized, universally web-accessible, secure "search" user interface (UI) was built as a single point-of-search entry for locating, viewing and/or downloading shared documents throughout the organization.

- Any search results generated by the system can be vertically filtered to narrow the results to include only those of a particular file type (e.g., PDF files) or part of specific collections. For example, if a user were to search for "immigrant eligibility," the search results can be further refined to narrow the results to only those documents part of LSNC's California Food Stamp Guide, or only those that are PowerPoint (PPT) files.

- Custom search result filters have been integrated into the search UI so that users with a single click can vertically filter search results by the following seven file types: WordPerfect (WPD), Word (DOC), Adobe Portable Document Format (PDF), Rich Text Format (RTF), Text (TXT), Excel Spreadsheet (XLS) or PowerPoint (PPT).

- Custom search result filters have also been integrated into the search UI so that users with a single click can vertically filter search results to narrow the search to specific special collections, including:

- o LSNC private intranet content
- o administrative forms
- o case forms
- o recommended exemplars for pleading and practice
- o comprehensive file collections for all impact cases and projects
- o LSNC public website content
- o California Food Stamp Guide

- Google "OneBox" modules — special GSA-driven database queries that return the data searched for as a special search result — have been integrated into the search UI. For example, the system offers universally searchable contact information for all LSNC personnel, by name or office location. If one were to search for "staff mark", those keywords automatically generate at the top of the search result page a list of all LSNC staff with the name "Mark" displaying their telephone numbers, email addresses, office locations, and photographs. Similarly, one can do a staff search by office location (e.g., keywords "staff redding") to generate a complete list of all LSNC employees in that particular office, with the same individualized information. (The Appendix includes a screenshot of how a OneBox module displays its special search results.)

- A comparable OneBox module has been integrated to return custom search results with links to all LSNC forms. For example, if a user were to search for "form reimburse," the GSA returns a special result at the top of the search results page with direct links to any and all LSNC forms that contain the word "reimburse."

- This same search UI was integrated with the SharePoint server to provide an intuitive, transparent method enabling staff to upload shared documents to any location within the structural taxonomy, and add select metadata as part of the upload process; the same process has been functionally integrated into "Save" file dialogs of all Microsoft applications in common use throughout the organization.

Search Targets

All authenticated users throughout LSNC can now search and/or navigate to all content that is targeted by this system. The system integrates all content in all designated LSNC private and public collections, including:

- all official LSNC intranet content (policies, protocols, forms, administrative and case manuals, personnel and other human resources information; development and fundraising resources; technology-related information, etc.) For example, users can quickly search for all LSNC forms or memoranda with the words "cash reimbursement," receive all but instantaneously a GSA-generated search result, from which one can click on the linked document to view or download it.

- a peer-reviewed collection of recommended exemplars of pleadings, memoranda and other types of commonly used advocacy documents. For example, one can search or navigate to these select pleadings for "demurrer unlawful detainer" to locate a recommended example of a memorandum of points and authorities in support of a demurrer in an unlawful detainer action.

- a comprehensive collection of supporting documents for all impact cases and projects. For example, one can search or navigate to this special collection to locate all final versions of all types of documents (pleadings, research memoranda, correspondence, etc.) created at all stages in all forums (trial, appeal, etc.) for dozens of major impact cases or projects.

- staff can also use the same single point-of-search to search for, locate and view or download all content at all LSNC public web locations, including:

    o the LSNC.net flagship site
    o the California Food Stamp Guide
    o LSNC's regularly published court case summaries
    o LSNC's regularly published housing and welfare regulation summaries
    o all of LSNC's publicly posted housing forms, fact sheets and other resources
    o all content at all of LSNC's five other special program websites, including its Health Rights Hotline; Health Insurance Counseling and Advocacy Program, Ombudsman Services of Northern California; Senior Legal Hotline; and Voluntary Legal Services Program

- hundreds of thousands of other user-generated, high-value character/text-based documents (WordPerfect, Word, Excel, PDF, PowerPoint, RTF and TXT files) to be organized taxonomically in shared document repositories.

Orientations, Trainings and User Evaluations

A combination of onsite and remote trainings were conducted for all LSNC staff program-wide to orient them to the new system, explain the UI and search results pages, and guide them through special search features and search filtering techniques. The initial trainings were conducted in October 2008 for the two test-bed locations, one onsite for the Sacramento office and one remotely for the Chico office.

With the benefit of having conducted those trainings, the initial survey results from the two test-bed locations, and informal feedback from a limited set of those same users, the project staff worked for the next several months, into the Spring 2009, on getting all GSA content targets identified, organized and in place; experimented with several iterations of the search UI ; incorporated custom filtering of search results; resolved the basic coding issues for creating GSA custom "OneBox modules" for generating search results from special data queries (as illustrated by the "staff" search described, above); and redesigned how search result pages display to make them easier to use.

At near completion of the project, an onsite training was conducted for the Redding office on March 25, 2009, to evaluate the latest changes made to the system. Three remote program-wide trainings were then conducted via LegalMeetings for all LSNC office locations, on May 26-28, 2009.

Online surveys were conducted at each of these major junctures. The surveys were conducted to evaluate user understanding of the new system, their success in searching for and locating what they were looking for, and whether the new system improved the efficiency, effectiveness and quality of their work. Their were 17 response to the first survey (of the Sacramento and Chico test-bed locations); six responses to the second survey (of the Redding office in March 2009); and 35 responses to the last, program-wide survey.

The survey results were overwhelming favorable. Among highlights of the survey findings were the following:

- 92% said they the organization (structural taxonomy) of the shared document repositories made sense (67%) or made perfect sense (25%).

- The overwhelming majority of survey respondents found the single point-of-search UI design understandable and organized effectively (67%).

- The basic Google search paradigm is extremely familiar to and well understood by all users: 100% of survey respondents understood how to conduct basic searches using already familiar Google search syntax. None reported any problem with conducting a basic search using the system.

- At the beta stage of the initial installations in Sacramento and Chico, 73% reported their searches as "successful," "very successful" or "dead-on successful."

- At project completion in May 2009, relying on the same descriptions of success used in earlier surveys, the overall "search success" rate program-wide increased to 90%.

- The understanding of search results paralleled the relative success of conducting the search: at the beta stage, 73% reported having a good understanding, a very good understanding or absolutely understanding the results obtained; at project completion this rate increased to 91%.

- At project completion, based on a four-star rating system, 89% of respondents rated the system at either 3 stars (39%) or four stars (50%) for its enabling them to more easily, quickly and efficiently find information within LSNC.

- Based on the same rating system, 82% of users rated the system at either 3 stars (36%) or four stars (46%) for its enabling them to be more effective at doing their work.

- Respondents overwhelmingly (71%) rated the system at either 3 stars (32%) or 4 stars (39%) for improving the overall quality of their work.

Post-project technical evaluation by Google Enterprise

Upon completion of the project, it was reviewed via a 90-minute remote session on June 22, 2009 by a four-member team of GSA specialists with Google's Enterprise Deployment Team. The purpose of the session was to conduct a technical and best-practices assessment, and offer follow-up recommendations.

During the session, the Google Enterprise team reviewed all basic system settings to determine whether they were active and properly configured, consistent with the project objectives. The review determined there were no technical or system problems or issues affecting the project as implemented.

Google Enterprise did recommend a number of "best practices" to encourage the project to exploit several configuration options not in current use, to further improve overall search results. These recommendations include implementation of "query expansion" so that the GSA returns results based not on just literal keywords but also words sharing the same word stem; use of "keymatch" features that will return a special search result at the top of the search result page (akin to a OneBox module) for specific keywords (for example, a search for "annual report" would generate a special, highlighted link to the organization's annual report at the top of the search results page); and use of special, GSA-specific report functions (akin to Google Analytics, but more specialized) for evaluating use of keywords, i.e., those that are most common, those that result in searches the user does not interact with, and so on.

The reviewing team specifically complemented the project on the design of its search result page UI design, for its visual clarity of where things are and what is available to the user.

Objective 4: Creation of "The Findability Project" development website, to promote and foster understanding and replication of cost-effective enterprise search solutions within the legal services community

One of the particularly innovative aspects of this project has been the creation of a public web development site, called The Findability Project, to offer transparency to the larger legal services community about the project, while facilitating understanding and replication of its goal and objectives. This public site was launched and promoted on July 21, 2008. Over the last 12 months of the project, it regularly published commentary, tutorials and analysis of the planning, technical details, and search concepts at the heart of the project. Among highlights of the project development website during this one-year period are:

- publication of 24 pages of tutorial content, including a 16-section guide on installing a configuring SharePoint as a component of an enterprise search platform

- 42 posts on varied enterprise search topics relevant to the project, including extensive analysis of search concepts, taxonomy, metadata and user interface design

- 600+ feed pulls per month

- 4,400+ visitor sessions

- 20,700 pages views

- an exceptionally low bounce-rate of 11%

- 15 published comments about the project, including comments from Patrick Lambe, author of [Organising Knowledge: Taxonomies, Knowledge and Organisation Effectiveness](); and [Daniel Tunkelang](), chief scientist at [Endeca](), a national, direct competitor of Google in the field of enterprise search.

- a request from the [Google Enterprise]() deployment team to profile [The Findability Project]() as an innovative, successful non-profit use of Google technologies.

V. Factors Affecting Project Accomplishments and Strategies Implemented to Address Major Challenges

Going into this project, LSNC had some practical advantages: LSNC had about five years of practical experience with Google's entry-level search appliance, the Google Mini, which it had long used for site search at LSNC.net. In more recent years, LSNC had also migrated toward use of the Google Site Search (a k a Google Custom Search) application, a virtualized site search platform now commonly used for public website search. Although the current enterprise-level Google Search Appliance (GSA) used for this project is a much more advanced, complicated, feature-capable device, the prior experience with the Google Mini and other basic Google search technologies like Google Site Search made it much easier for LSNC to understand and adapt these same technologies to our organization's enterprise search objectives. Quite literally, the day the Google Search Appliance was delivered, we were able to connect it to a test target server and initiate a successful, if rudely elementary, crawl and return basic, unrefined search results.

LSNC's IT staff had no significant problems in building out the hardware and software infrastructure required for the project. Purchase of the GSA also included two years of direct GSA device technical support from Google Enterprise. That said, going in we also knew the scale and sophisticated technical demands of a more advanced Google Search Appliance would require reliance on a professional GSA consultant to achieve our larger

project goal.

After reviewing our project goals and objectives, Google Enterprise recommended [Michael Cizmar + Associates](#) (MC+A), a Chicago-based information system design firm prominent in the GSA community. MC+A was extremely helpful, if anything all but indispensable to the project as we set up our GSA configuration, established all our enterprise search targets, coded our custom point-of-search UI, tested and improved our search results, and created special custom filters for refining search results. There were also a number of incidental but not inconsequential challenges in properly configuring and reconfiguring the open source Google connector to our SharePoint server installation, which MC+A quickly resolved for us remotely.

There was one major technical challenge presented by this project that we were unable to resolve. LSNC discovered that there are significant, unanticipated technical challenges unique to the Pika CMS that thus far have prevented effective use of the GSA to target Pika content. The problem is not "targeting" Pika with the GSA, which by design performs wholly benign, non-destructive crawls as it indexes targeted records. Rather, the problem is that the current version of Pika is not well optimized for use as an enterprise search target. There are code anomalies in Pika that, among other things, cause it to auto-generate new case intakes and case records when it is crawled by the GSA. Further, there is the need to optimize Pika code to create within the application more index-friendly listings of case records, notes, documents and user accounts that can be effectively exploited by the GSA. The project presented these Pika compatibility issues to its developer, Aaron Worley, who has analyzed the problems and recognizes the need to retool Pika to make it more suitable as an enterprise search target by LSNC and other legal services programs implementing Google Enterprise and comparable search systems. At this juncture it is uncertain when Pika Software would expect to fix the code anomalies creating the problem.

Even in the absence of a current solution for integrating Pika search results directly into our larger enterprise search solution, LSNC staff are able to rely on Pika's native search functions to locate all essential client case record information, and then rely on the GSA to locate everything else within the LSNC search domain.

LSNC is very committed to finding a solution to this problem. Even though it is now clear that a conventional GSA crawl will not work properly with the Pika CMS, we will be exploring other approaches, including techniques whereby the GSA culls data directly from the MySQL database used with Pika, outputs that data as XML, and then independently of Pika publishes the XML output to an XSLT template suitable for crawling and indexing by the GSA. Should those efforts be successful, details about the solution will be posted to The Findability Project.

VI. Major Lessons and Recommendations

Lessons Learned

- **Process and transparency are vital to the success of a project of this scale and depth of impact on an organization.**

Making sure everyone "gets the memo" just does not cut it. It is perhaps too obvious a management truism to state the above heading, but it is one that bears repeating in a project like this one. It would not have been enough to announce that the organization would create new search features and that they would work like Google and, well, leave it at that.

This project directly implicated how all staff work, what documents they would openly share with others to better perform their work, how easy or hard or convenient or inconvenient it would be to locate and retrieve documents in the system, how all the documents and files constituting so-called "knowledge content" would be identified and then organized, and how staff going forward would contribute to that shared content. There are no short cuts to communicating with staff and making it transparent what the project is about and why it is important and will be helpful to them in doing their work on behalf of their clients. The process of vetting major project elements, design and implementation, as well as vetting proposed practices, protocols and best practices, were essential to making a project like this work for all involved. Talking and listening and vetting matter, and make a difference, and it should be apparent to your staff what that difference is.

- **It is critical to a project of this scale and complexity to retain a professional consultant who has real-world expertise in information architecture and related systems, expertise in implementing your chosen enterprise search platform, and can realistically assesses the organization's needs and resources for successful completion of the project.**

We were able to advance this project fairly efficiently because we spent several months seeking recommendations for a GSA consultant and actually conducting interviews with three of them. It helped immeasurably the Google Enterprise itself, after discussing our project with us, provided us with specific recommendations, all of which we pursued. That is how we found our consultant, MC+A.

After working with our consultant, we experienced the dividends of doing our ground work on identifying a consultant with a good fit for our project. Good consultant do not come free, and we budgeted this project accordingly. That said, we got every dollar's worth out of the arrangement. Our consultant MC+A consistently demonstrated a high level of GSA expertise, while always sensitive to the limited resources we could bring to this project. The consultant made sound, affordable recommendations at every step of the way, whether it was recommending open source technologies like the GSA SharePoint connector over a proprietary connector, or other low-cost alternative like the

open source XSLT and CSS coding we ended up using, or suggesting ways we could simply our taxonomy in a practical ways that would make the directory structures more useful for staff without compromising the quality of search results. There were innumerable technical and practical improvements we gained from reliance on a qualified GSA consultant.

- **The Google Search Appliance offers advantages over other enterprise search platforms because it seamlessly integrates with Google Apps and other Google tools commonly used in legal services and other non-profit organizations.**

At the outset of this project, we instinctively understood that there were likely some practical advantages we did not yet appreciate about the Google Search Appliance because it was a Google product, and as an organization we were already using many other Google products. As we got deeper into the project, it became self-evident that we could quickly integrate our organization's use of various Google applications – most notably Google Sites for our organization's cloud-based private intranet – as GSA targets. That realization prompted our decision to decommission our existing intranet content at a old MediaWiki site, abandon plans to use the SharePoint web publication features, and rebuild and reorganize all our intranet content using Google Sites. Currently we are able to use the GSA to target and make easily "findable" selected content in our domain's Google Sites and Google Docs. Plus, Google Analytics works natively from within the GSA itself and returns very useful, easily understood metrics for analyzing searches results and search behavior. These are excellent bonuses that come with adoption of the GSA platform.

Recommendations

Creation of a "knowledge content" system and adoption of suitable "document management" tools are among the baseline technology recommendations issued by LSC in May 2008. It is all but inevitable that, at some juncture, all legal services field programs will implement purposeful document management systems and will adopt some type of enterprise search solution so that all shared content becomes easily and intuitively "findable" to all in the organization.

There is little purpose to having a knowledge content system, however, if it is not sufficiently comprehensive, not well organized, and the content is not findable. Based on our experience with the development and implementation of this project, we strongly recommend that other legal services programs seriously assess their need for implementing a comprehensive knowledge content system that, in a cost-effective way, integrates enterprise search. The improvements in the "findability" of the organization's shared knowledge content, and how that findability enhances the efficiency and effectiveness of the work done on behalf of the clients it serves, can be very dramatic.

Designing and implementing such systems are anything but "turn key" operations. But there are a range of ways legal services program can begin to build in-house experience and expertise with search technologies, even if current fiscal circumstances limit their

options at the moment. Using the Google options familiar to us, we can recommend field program consider implementing:

- Google Desktop and/or Google Desktop Enterprise, both cost-free options providing individual and local office search capabilities from the desktop and through drives mapped to a local network.

- Migrate to Google Apps so that the program can offer cloud-based intranet features throughout the organization, including Google Docs and Google Sites, both of which have very effective, basic Google search functionality.

- Step up to an entry-level search appliance with the Google Mini, which can handle from 50,000 to 300,000 files. The Google Mini is not as fully featured, nor does it offer the expanded document capacity, as the full-on Google Search Appliance used for this project. But the Google Mini shares all the basic functionality and custom search result features and works with both public web and private internet content. (An important *caveat*: At the time of this report, the Google Mini does not integrate with the Google Apps platform.)

These options will be explored in an upcoming LSNC presentation sponsored by NTAP in July 2009, entitled "Findability and the Google Search Paradigm: Integrating Search as an Organizational Solution." They will also be discussed as part of another LSNC presentation in January 2010, when it conducts a similar, more substantial, technological session about The Findability Project at the upcoming 2010 TIG conference.