



A preview of Lotus Discovery Server 2.0

Level: All
Works with: Discovery Server 2.0
Updated: 01-May-2002

by Wendi Pohs
and [Dick McCarrick](#)

Knowledge Management (KM) is a term that's been getting quite a workout the last year or two. On one level, the words are almost self-explanatory. What organization wouldn't benefit from better managing its collective knowledge? But many people still have questions about KM. How exactly does it work? What KM products does Lotus offer? And how can they help my organization?

This article will help answer these questions. We'll start by briefly examining KM, what it means and how it can benefit you. We'll then review the Lotus Discovery Server, Lotus's KM server offering. Finally, we'll discuss feedback we've received from our Discovery Server customers, and how we've responded to this feedback in Release 2.0.

Knowledge Management: Well, what do you know?

Few would dispute that knowledge ranks among an organization's most precious assets—if you can find it when you need it. But knowledge is not always easy to locate. Ideally, combined intellect and experience of the entire group should be readily available to everyone all the time.

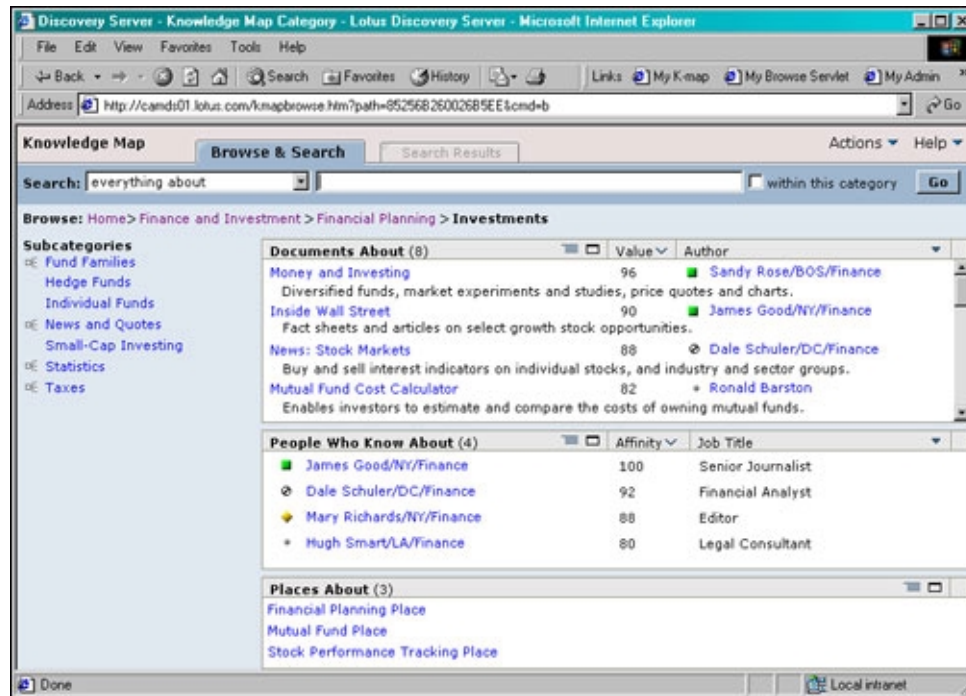
With some of the tools available today, you can search for all the information your organization or the Internet has collected on a topic, but too often the results are disappointing. The information is too voluminous, unfocused, or out of context—or simply too difficult for the average person to understand. What users really need is experience—the perspective of someone who has done it before. That experience could come in the form of a document, or (even better) a person to whom you can look for assistance. So what's needed is a product that speeds the knowledge management process in a way that brings new scalability to the task of analyzing your ever-growing volume of content. It should also eliminate the need to manually process and manage knowledge. The Lotus Discovery Server answers these needs.

The Lotus Discovery Server primer

The Lotus Discovery Server is a back-end server for managing your organization's knowledge. The Discovery Server provides sophisticated tools that categorize documents and user information into browsable and searchable form. These tools include the following.

Knowledge Map, or K-map

The Knowledge Map or K-map (also called the taxonomy or catalog) is a graphical representation of your organization's knowledge. It displays a hierarchical set of categories and documents you can use to find information. The K-map is the backbone of the Discovery Server search-and-browse user interface. From the K-map interface, you can locate content from many disparate sources, by drilling down through subject categories, using full-text search, or using a combination of both search strategies. Additional information about the relationships between people and document activity adds value and context to the user's search and retrieval experience. Because the K-map displays related documents, people, and places in categories, users can browse and search for information in context.



K-map Building Service

The K-map Building Service creates the K-map, which you can subsequently modify using the K-map Editor (explained next). The K-map Building Service builds document categories, creates labels for these categories, and places new documents into existing categories. It also identifies documents that do not fit into any existing categories.

K-map Editor

The K-map Editor is a client application that lets you fine-tune the K-map to meet the needs of your organization. Neither the K-map Building Service nor any other automatic process can predict precisely how an organization wants to structure its content. It can only build a K-map based on the words in the content. Once the basic K-map is built, you can use the K-map Editor to drag categories from one level to the next, re-label them with preferred terms, and place documents in different categories. This in turn helps teach Discovery Server how to categorize documents with similar content in the future.

Profiles

[Profiles](#) help identify the right people for the right job. Profiles collect existing user information from the directory and other sources, providing a more complete representation of the users in your organization.

Spiders

Spiders are multi-threaded processes that collect data. This data can exist in a number of different file formats, including:

- XML
- Exchange e-mail and public folders
- Web content
- Windows-compatible operating system files
- QuickPlace
- Domino.Doc
- Notes databases and e-mail

Once the spiders collect this data, the K-map Building Service processes it to create the K-map.

Metrics

Metrics, which is also called affinities processing, is a computational program that looks at existing documents and relationships between documents and people. The metrics component does two things. First, it calculates the value of a document. Second, it calculates an affinity between a person and categories, based on the person's

interactions with documents in the categories, which in turn helps produce category affinities.

Other tools

Additionally, administration tools let you install, set up, and maintain Discovery Server, and security features protect your data.

For more information about KM and the Lotus Discovery Server, see *Practical Knowledge Management: The Lotus Knowledge Discovery System* by Wendi Pohs, which is available from [IBM Press](#).

Customer experiences and feedback

Lotus launched Discovery Server 1.0 in early 2001. Companies adopted Discovery Server to help manage their corporate knowledge, especially when they had an immediate informational problem to solve. But as with any new product in a leading-edge technology, there was a lot to be learned as organizations began to implement Discovery Server and use it in many different ways—some expected, some unexpected.

For example, we initially assumed that once users manually categorize documents, they would want these categorizations to "stick" and not be further modified by Discovery Server. So when automatically updating a K-map, the K-map Building Service ignored all documents that had been manually categorized. However, we soon learned users wanted help in creating subcategories, especially when a category grew so large that it became difficult to use. So in Discovery Server 2.0, the K-map Building Service now processes manually categorized documents.

Customers demanded a more open, less "black box" approach. In response, we're providing more detailed administration functionality and interface. We're also opening the Discovery Server API, to allow more customization. Users also want the ability to import an existing operating system file/folder taxonomy into Discovery Server and display it as a K-map.

Early Discovery Server adopters also requested improved metrics reporting (particularly for information trends and knowledge gap analysis), better automatic profile creation (to avoid re-keying data that already exists, for example in a Person document in the Domino Directory), and better metadata handling (for instance, more accurate connections between documents and their authors). In Discovery Server 2.0, we've attempted to respond to all key points in this feedback.

What's new in Discovery Server 2.0?

So what's new in Lotus Discovery Server 2.0? The remainder of this article discusses the new features and functionality we'll be introducing in Lotus Discovery Server 2.0 Highlights include:

- A revamped architecture for the K-map user interface and editor, with more accessibility and better display of search results
- Improved "people awareness" for Profiles
- Better spider support for Domino.Doc and QuickPlace content
- Enhanced control over metrics processing
- Easier installation, setup, and maintenance
- Better logging

K-map: Improved interface, expanded features

The 2.0 K-map offers a user interface that features a new servlet-based architecture. This gives you more scalability and better performance. Other K-map enhancements include accessibility, improved display of search results, bookmarking, and better support for opening the most appropriate replica of Notes databases.

Easier-to-use, more accessible interface

The new K-map user interface consists of two main tabs, Browse and Search, and Search Results. In addition, a new URL-addressable page appears for all the following actions:

- Browse to a new category
- Perform a search
- Navigate to a different tab within the same set of search results
- Perform a refined search
- Navigate between sets of iterated search results
- Page through a set of list results
- Sort a list

Each page that appears as a result of these actions is put on the browser history, so you can return to it via the Back/Forward buttons in the browser interface. You can also bookmark these pages.

Lotus is also committed to offering an accessible interface for Discovery Server. This includes supporting all functionality with consistent and simple keyboard navigation of Discovery Server controls, fields, and hyperlinks. We will also aid low-vision users by:

- Providing enhanced color contrast in the standard interface design
- Supporting the Windows "High Contrast" palette schemes in the Appearance tab of the Control Panel - Display settings
- Supporting user-defined font settings in both the operating system and the browser

Other K-map user interface features introduced in 2.0 include:

- K-map search results offer People summaries in search results. We've also added a "Go to K-map Category" link to the summary of each document in search results.
- Supported documents types include documents stored in Exchange public folders. Discovery Server 2.0 also offers better handling of Notes attachments, OLE embeddings, Domino.Doc documents and QuickPlace documents.
- Bookmarking lets you bookmark (whether from the Search or Browse page) and have Discovery Server save both the Browse state and the Search parameters.
- Saving state maintains a single saved state so that whenever you access the K-map from a particular computer, the interface is configured exactly as it was the last time you logged in. In response to customer feedback, the K-map no longer returns you to the category you were in when you last closed the K-map. This helps improve performance. Key information saved through this feature includes:
 - Document and People List heights in Browse and Search
 - Column order, in Document and People Lists in Browse and Search; and on Document, People, and Category tabs in Search Results
 - Sort order in Document and People Lists in Browse and Search
 - Column widths, in Document and People Lists in Browse and Search; and on Document, People, and Category tabs in Search Results
 - Summaries on/off, in Document and People Lists in Browse and Search; and on Document, People, and Category tabs in Search Results

K-map Editor

The K-map Editor interface now lets you display a special K-map Editor accessible Reports view. This lets you create, schedule, and delete two new reports to aid your editing. These reports show both the number of documents per category, subtotaled per branch, and the documents that are new to a category (having been added automatically or manually).

Another new K-map Editor feature is category visibility. This option allows you to keep categories hidden until they are ready to be viewed by your users. Then when you publish a category, it becomes visible to end users via the K-map interface, as long as its parent categories are also published. (New categories assume the visibility of their parent by default.) Note that hiding a category does not prevent users from seeing the documents in that category if they are returned via K-map search. However, it does prevent you from getting affinities to that category.

To make it easier to optimize taxonomies, Discovery Server 2.0 offers two new options, Request Subdivide and Request Retrain.

Request Subdivide tells the K-map Building Service to divide the selected category into subcategories. When you select this option, the K-map Building Service attempts to create these subcategories, based on the total number of documents in the category (and the maximum number of documents per category you specified). If the K-map Building Service determines it can create two or more subcategories with reasonable fit values (and fairly evenly distributed documents), it moves documents from the specified category to the new subcategories, leaving the original category empty. If the selected category can't be usefully subcategorized, the K-map Building Service returns an error message.

Request Retrain helps "teach" the K-map Building Service how to categorize documents the way you want. After you manually move documents into a category, this option tells the K-map Building Service to place new documents with similar content into this category in the future. You can retrain selected categories, or the entire K-map taxonomy.

Other K-map Editor enhancements include an icon that indicates whether documents will launch in their native application or in the browser, a set Document Status options to set status for selected documents to locked or unlocked, and a Doc Counts option that displays the number of documents per category

K-map Building Service

In response to user demand, we've made the K-map Building Service smarter when handling categorization. For example, in Discovery Server 1.0, if you manually moved a document into a category, K-map Building Service automatically assumed you wanted it to remain there forever, even if the category grew so big it required subcategorization to navigate properly. It would never change the categorization of this document. In Discovery Server 2.0, the K-map Building Service is free to subcategorize these documents as appropriate.

By default, all categories created by the K-map Building Service are hidden, so you can review them prior to making the categories available to users.

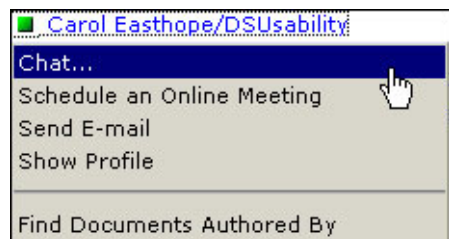
You can also import an existing file system taxonomy and use it as the basis for your K-map. To do this, Discovery Server 2.0 lets you import the file/folder document taxonomy from your operating system. You can then use the K-map Editor to modify the taxonomy into the K-map you want.

Profiles: Who's available now

Profiles include enhanced people awareness, the ability to determine the on-line status of selected members of your organization. We've also upgraded Person profile documents, which now include a more accessible interface.

People awareness

People awareness incorporates Sametime functionality to transform passive name references into dynamic resources. These provide information about the person's current on-line status. Directly from the name reference, you can contact the person, find out more about them, and initiate other application-specific commands.



When you log on Discovery Server 2.0 (either via the K-map or through a profile), you automatically log on to the Sametime server. Then anywhere your name appears within Discovery Server, other users will be able to tell you are on-line. Additionally, wherever they see your name displayed in K-map or in a profile, a Sametime status icon appears to indicate your on-line status. (If you don't have a Sametime server specified in your Domino Directory Person document, Discovery Server assumes Sametime is not available and doesn't display your on-line status.)

People awareness lets you identify your most knowledgeable people in a particular area, and then contact them immediately by initiating a Sametime dialog. This is especially useful when you need immediate information or need a question answered in real-time.

Person profiles

We've given Person profile documents a new servlet-based interface, to ensure the same "look-and-feel" as K-map. This also gives you better support for text resizing using your browser's View - Text Size commands. We've also made the Person profile interface more accessible.

One Discovery Server customer request is the ability to see all names the user is known by throughout the system. So at the bottom of the Contact Information page, we've added a new field called "Other user names."

We've also made some changes to the Affinities interface. We've moved the interface for approving proposed affinities into the profile document. For example, if there are proposed affinities, they appear in a table within the profile document. (Note that if you approve proposed affinities but then cancel out of the profile document without saving it, the approvals are also cancelled.)

And in response to other user feedback, we display the "Declaring Affinities" description in Read Mode on the Affinities page if the person viewing the profile document has edit rights.

Spiders: Better support for Domino.Doc and QuickPlace

Discovery Server 2.0 supports spidering Domino.Doc and QuickPlace documents. We've also enhanced other spiders, such as the one for Notes/Domino, and added a spider for Exchange.

Domino.Doc spider

We offer spidering for both Domino.Doc 3.1 and 3.0. Capabilities include:

- Documents and their attachments are treated as single documents.
- A Domino.Doc repository is defined at the File Cabinet level.
- Forum (discussion) docs are spidered.
- There is an administration option to spider all, or only latest versions.
- Domino.Doc-specific field mapping is spidered.
- There is support for archived documents.

QuickPlace spider

We also support spidering for QuickPlace 2.0.6a and 2.0.8. Capabilities include:

- A QuickPlace repository defined at the Main.nsf level.
- Embedded pages are treated as one entity.
- QuickPlace-specific field mapping is spidered.

Handling documents with attachments

Spiders in Discovery Server 2.0 have improved the handling of documents that do not contain text but do contain one or more attachments. These so-called "sparse container" documents will be classified with their attachments. The title of the attachment identifies the sparse container document.

Metrics: More detailed and easier to understand

Metrics now consists of four separate services: Profile Maintenance, Metrics Reporting, Affinity Processing, and Metrics Processing. These comprised a single service in Discovery Server 1.0; we separated them in response to user demand for a less "black box" approach to Metrics. All four services can only run on one server at a time; but you can move them from one server to another, and schedule them individually.

- Profile Maintenance processes user edits to profile documents.
- Metrics Reporting creates Metrics reports.
- Metrics Processing computes document values, updates the Discovery Server data with new affinities and new document values, and updates full-text search with new values.
- Affinity Processing calculates affinities, proposes and publishes affinities, sends affinity e-mails, and updates profile documents with affinities. This service provides more control over when affinities are generated.

Installation, setup, and maintenance: More customer control

Discovery Server 2.0 offers significantly enhanced administration functionality. This lets you have greater control over installation, setup, and maintenance. And you'll have a better idea of what's going on "under the hood" of Discovery Server.

Installation

The installation dialog box now includes numerous checks and warnings to better guide you through installing and upgrading. Also, we no longer install the K-map Editor with the server. We have discovered that in practice, installing the K-map Editor on the server is rarely done—usually only for demo purposes.

We also modified the Admin Name & Password screen to let you provide your own username. This can be an account already defined on the system or created on the fly if one doesn't exist.

Setup

In Release 2.0, we re-worked the setup screens to be easier and more intuitive. This includes better input validation and error messages.

Maintenance

As a response to customer requests for more control over Discovery Server (and to accommodate new features introduced in Release 2.0), we have incorporated new administration functionality. For example, we made the interface for enabling the XML spider always visible, instead of requiring you to set an INI variable to do this.

We've also added a new interface for replicating K-map data to a secondary server. This includes a "K-map Replication" checkbox in the Server document on the primary Discovery Server, as well as a "K-map Replica" checkbox in the Server document of Secondary Discovery Servers. After you check this option and save the Server document, replication will copy the K-map data from the primary to the specified secondary servers. These secondaries can then serve end-users, helping you balance user workload among several machines.

Other maintenance features include:

- A new interface for enabling the Exchange spider
- A drill-down model to support large numbers of repositories
- Better "paging" capability in multi-page views and logs
- The ability to move Metrics processing from one server to another
- Improved interface for choosing repositories for creating the K-map
- Updates to Service Status views

Logging

We've improved logging messages to be more informative. For example, the K-Map Building log has been significantly enhanced, with more logging of editor activities. And to keep you better apprised as QuickPlace and Domino.Doc spiders run, we provide a new message type to enable these spiders to post interim begin/end messages as each room/binder is processed.

Other enhancements

Discovery Server 2.0 API Toolkit

As mentioned earlier, users have asked for a more open approach. To meet this need, we're providing a complete Discovery Server API Toolkit. This allows you to customize Discovery Server to suit your organization's exact requirements. Third-party software developers will also use this Toolkit to develop their own solutions based on Discovery Server functionality.

Data repositories

Discovery Server 2.0 includes many new features to better manage your data repositories. For example, you can temporarily prevent a repository from being spidered, and have it start up again later, without disabling the spider. You can then have the repository requeued automatically. You can also stop spidering a repository (for instance, because you made a mistake in defining it), delete it, and start over with new parameters. Other new data repository functionality include:

- Preventing two repository records from pointing to the same source
- Supporting a "delete and make new copy" action from the main Repositories view
- Providing the ability to unqueue a repository
- An interface for spidering Exchange e-mail and Public Folders
- Options for traversing Domino.Doc and QuickPlace hierarchy
- Options for spidering Domino.Doc and QuickPlace revisions
- The ability to edit Field Map fields after the repository has been queued/spidered

And we've added an interface for specifying a subset of repository data to be spidered, per customer request.

Field mapping

Our goal in this release is to only map what the administrator selects to map—in other words, what you define in the profile forms. This means there will be no "identity mapping" (automatically mapping a field name in the source to the same field name in the profile document). Also in Release 2.0, we provide better coverage for well-known data mappings in the default data field map (\$Global).

People sources

New features in this area include

- An interface to specify "Process all documents during next run"
- An interface to associate supplemental sources with all or some subset of authoritative sources
- Support for additional LDAP search parameters BASEDN and SCOPE

Documentation

Discovery Server 2.0 documentation includes a Deployment Guide, available shortly after product ship. This should be a must-read before a customer site begins to install and implement Discovery Server.

Going forward

A director of research at a major automotive firm recently stated, "If we just knew what we know, this organization would be 30 percent more profitable." Lotus Discovery Server helps you move forward toward such goals. Through the years, members of your organization have conducted business, written documents, contributed to online discussion forums. In the process, they've created a collective body of information of ever-increasing value. Discovery Server 2.0 helps you make sense of it all. Less time is wasted spent looking for the person who knows about a particular subject, or locating the information you need. Your best assets are accessible to all, enabling employees to better meet the demands and expectations of their jobs. This in turn helps form high-performing teams—and your corporation performs better as a whole.

ABOUT THE AUTHOR

Wendi Pohs is a principal taxonomy specialist on the Discovery Server team and the author of a book about knowledge management methodologies, *Practical Knowledge Management: The Lotus Knowledge Discovery System*, published by IBM Press. Wendi joined Lotus Development Corporation in 1996 and has worked on various projects as a spec writer, online help designer, and user assistance manager. Prior to joining Lotus, Wendi worked at the American Mathematical Society and at Digital Equipment Corporation. Wendi received her BA and MILS degrees from the University of Michigan.