

Level: All
Works with: Discovery Server
Updated: 03-Feb-2003

by
Jim
Riel

With its release of 2.0.1, Discovery Server fulfills the promise of release 2.0 and responds to customer requests to give them a deeper level of control of their implementation. Discovery Server 2.0.1 includes an alternative, accessibility-enabled user interface that makes it possible to use all K-map features using a screen-reader. Besides accommodating users with physical challenges, the accessibility-enabled K-map can be accessed by versions of Internet Explorer prior to version 5.5 or by Netscape Navigator or Mozilla.

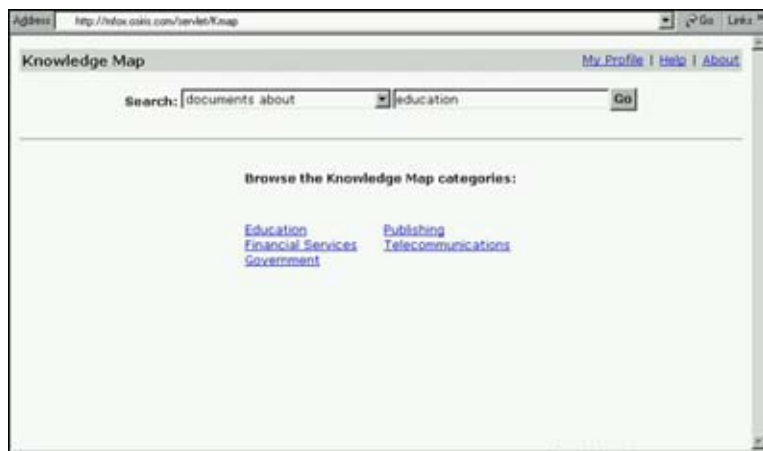
In addition, new administrative features provide the ability to:

- Specify a spider run as Preview Only to check the meta-data content of a repository before it gets committed to K-map creation.
- Override attachment meta-data with container meta-data in situations where the attachment's meta-data is sparse or inaccurate.
- Specify a repository for search-only.
- Configure a search-only Discovery Server, an implementation that doesn't rely on the K-map.
- Generate affinities only for a subset (group) of users at a time.
- Re-generate affinities (including deleting existing approved affinities from profiles).

This article describes these features in detail. We assume you're familiar with Discovery Server and its basic functions and terminology. For more information on Discovery Server 2.0 features, see the *LDD Today* article, "[A preview of Discovery Server 2.0](#)."

Accessibility-enabled K-map

Discovery Server 2.0.1 delivers a servlet application designed to accommodate people who have physical challenges, such as restricted mobility or limited vision. This servlet application makes it possible to use all K-map features via a screen-reader instead of the standard graphical interface (see the following screen).



Notable differences from the graphical K-map include:

- There are no tabs for Browse & Search versus Search Results pages.
- There are no menus (specific actions are called out individually on the Action bar as links).
- The Search user interface is centered on the page.
- Top-level categories are centered on the page and fill two columns.
- There is no subcategory icon.
- Background color is white, while links are the default browser style (typically blue and underlined).
- There is no image on the home page.
- Single Sign-on (SSO) works from the end user's perspective just as it does in the standard graphical K-map.

Although provided primarily to enable persons with disabilities to have access to all Discovery Server features, the accessibility-enabled K-map servlet also makes it possible to use versions of Internet Explorer prior to version 5.5 or to use Netscape Navigator or Mozilla. Instead of displaying an error message when users connect to the K-map using these browsers, Discovery Server automatically defaults to the screen-reader accessible K-map user interface. Also, we recommend that all applications that integrate with Discovery Server (such as Atomica and WebSphere Portal Server) use the URL to the accessible K-map.

Users with Internet Explorer 5.5 or higher who want to connect to the screen-reader accessible K-map must type in an alternate URL:

<http://servername.domainname.com/servlet/Kmap>

Alternatively, an administrator can make connecting to the screen-reader accessible K-map more automatic for his/her end users. Refer to the [Discovery Control Center administrative help](#) for more complete instructions.

The People Awareness component of Discovery Server is also fully featured except that:

- The appearance of people names only appear in default link style (which typically means underlined).
- A user cannot select multiple names.

Refer to the K-map online help for detailed descriptions of the differences between the graphical K-map and the screen-reader accessible K-map.

Previewing repository meta-data

For Discovery Server, the information contained in a data repository is only as good as its meta-data. This includes author information, keywords, document titles, and attachments. Meta-data helps Discovery Server to spider data repositories efficiently and to categorize and add documents to the organizational K-map. The more refined the meta-data, the better Discovery Server is at mapping the information. In order for Discovery Server to display repository documents effectively, and in order for the Metrics Processing service to gather accurate information, it is important that repository meta-data be as rich and complete as possible.

Discovery Server 2.0.1 includes a new feature that enables you to preview the state of a repository's meta-data prior to committing that data to the taxonomy by spidering the repository and generating a report about the documents contained in the repository. The preview report is an XML document that provides statistical information about the repository itself—such as the total number of documents and the percentage of documents

that lack a title—as well as each of the documents found in the repository.

Previewing this information enables you to better assess how valuable the repository might be to K-map creation and most importantly, to make a decision (before spidering the repository) as to whether or not you need to fix problems with the repository's meta-data. For example, if the titles and/or authors of the attachments are inaccurate, you can choose to use the meta-data from the container documents of the attachments instead. Or perhaps you make sure that all documents within a repository have accurate author information before the repository is spidered.

You can preview all repositories—including Notes databases, Web sites, XML files, Microsoft Exchange, and file systems—except email databases.

Generating the preview report

The preview report is output to a Preview directory in the Discover Server data directory specified in the Notes.ini file (usually \Lotus\DS\Data). The report file name is the name of the repository. An example of a path name and file name for a preview report is:

C:\Lotus\DS\Data\Preview\RepTitle.xml

You set up a preview using Discovery Server's Control Center. The first step is to add a repository:

1. In the Data Repository view, click Add Repository. This opens a data repository form that you complete with repository information.
2. You set up a preview feature in the Spider Schedule section. Select Yes for Schedule Enabled.
3. In One-Time Settings, select the option "Only generate preview file of meta-data during next run."

4. Save your changes. The repository will be spidered for the preview report the next time the spider runs. By default, the preview option is not enabled, so you must set up the preview option intentionally.

Viewing the preview report

To view the preview report, open the C:\Lotus\DS\Data\Preview file folder either in Windows Explorer or by using My Computer. Double-click the name or icon of the Preview report you want to see. The first part of the report contains information about the repository in general and looks similar to this:

```
<?xml version="1.0" encoding="utf-8" standalone="no" ?>
- <repository title="Fixumup">
+ <statistics>
  <total_documents>4631</total_documents>
  <total_missing_title>0</total_missing_title>
  <percent_missing_title>0.00</percent_missing_title>
  <total_missing_author>0</total_missing_author>
  <percent_missing_author>0.00</percent_missing_author>
  <number_attachments>6</number_attachments>
  <attachments_missing_title>0</attachments_missing_title>
  <percent_attachments_missing_title>0.00
</percent_attachments_missing_title>
  <attachments_missing_author>0</attachments_missing_author>
```

```
<percent_attachments_missing_author>0.00
</percent_attachments_missing_author>
</statistics>
```

The rest of the preview report data contains information about specific documents in the repository. For a Notes database, the information looks something like this:

```
- <attachment>
  <title>Designware Software Demo(attachment - myhdemo.bat)</title>
  <author>Mel Dunphy</author>
  <httpurl>HTTP://mfox.osiris.com/FIXUMUP.NSF/0/BF4B8B74A870A045852561640076F7F0?Open</httpurl>
  <nativeurl>NOTES:///8524748D0075D914/0/BF4B8B74A870A045852561640076F7F0?Open
  </nativeurl>
</attachment>
- <attachment>
  <title>Designware Software Demo(attachment - install.olb)</title>
  <author>Mel Dunphy</author>
  <httpurl>HTTP://mfox.osiris.com/FIXUMUP.NSF/0/BF4B8B74A870A045474861640076F7F0?Open
  </httpurl>
  <nativeurl>NOTES:///8524748D0075D914/0/BF4B8B74A870A045852561640076F7F0?Open
  </nativeurl>
</attachment>
<document>
  <title>building supplies needed</title>
  <author>Ken McGiver</author>
  <httpurl>HTTP://mfox.osiris.com/FIXUMUP.NSF/0/000003AB86611B4547485D1000714D7C?Open</httpurl>
  <nativeurl>NOTES:///8525696D0075D914/0/000003AB86611B4547485D1000714D7C?Open</nativeurl>
</document>
- <document>
  <title>I have some left-over materials from an addition</title>
  <author>Diane Lane</author>
  <httpurl>HTTP://mfox.osiris.com/FIXUMUP.NSF/0/000001ED11B464748255D11004EF817?Open</httpurl>
  <nativeurl>NOTES:///8524748D0075D914/0/000001ED11B4610285255D11004EF817?Open</nativeurl>
</document>
```

Spidering a repository for search results only

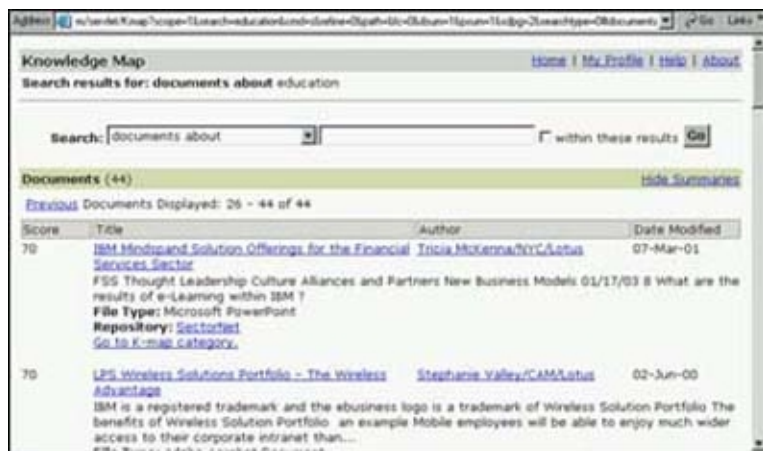
There may also be situations in which you want information about the contents of a repository before you commit it to the K-map. Perhaps you're not sure what's inside, and you want to determine if the contents belong in your taxonomy (and if so, where). Or perhaps you want to provide your users with the option of searching a repository without having to use the K-map. In Discovery Server 2.0.1, you have the option of spidering a repository for search results only (see the following section). If this option is enabled, the repository is spidered on a scheduled basis, but the documents from the repository are not sent to the K-map Building queue and are not included in the K-map categories. You might choose the search-only option if you want users to get to your data quickly or if it's data that's worth retrieving, but maybe not robust enough to use to generate affinities. Heavily fielded data repositories are great candidates for search-only.

Setting up search-only

You enable the search-only option when you create a new Data Repository definition form. The option to spider the repository for both K-map categories and search results is enabled by default. To enable search-only, click K-map search results only before running the spider. Note that if you are upgrading from Discovery Server 2.0, existing Data Repository definition forms are automatically set to the option "Both K-map categories and search results."

1. Open a new Data Repository definition form.
2. Specify the location parameters as you would for any data repository plus any detail specific to the repository type.
3. Instead of selecting the option "Both K-map categories and search results," choose K-map search results only.
4. Schedule the repository for spidering as you would any other repository, then save the form.

Once the spider has been run, users can access the repository's contents using the Knowledge Map's Search toolbar:



You can also create a customized search user interface for users. See the Discovery Server API Toolkit Help for more information on building a customized search user interface.

After the spider for the repository has been run, you cannot disable the search-only option. This means that once the repository has been spidered for search-only, it will always be spidered for search-only. If you later decide that you want to include the documents from this repository in the K-map, you must delete the repository's definition form, then re-specify it, this time selecting the option "Both K-map categories and search results." You can use the Delete but reuse definition action to re-specify the repository more quickly because this action retains most of the parameters previously specified.

If a repository has been spidered for both K-map categories and search, you can also use the search-only option to remove repository documents from K-map categories, but leave them available for search. To do this, you must delete the repository's definition form and re-specify it, this time selecting the option K-map search results only. Again, you would use the Delete but reuse definition action to re-specify the repository more quickly. When the repository is spidered for search-only, the repository documents are removed from the K-map.

Repositories that have been enabled for search-only do not appear in the Spidered Repositories by Type view when you're specifying repositories to use for creating your K-map. This view displays only those repositories (enabled and disabled) that have been enabled for "Both K-map categories and search results" and that have been spidered already.

Implementing a search-only Discovery Server solution

Data repositories designated as search-only, as described previously, can fit easily into a full Discovery Server implementation that utilizes a K-map, where you want the search-only repositories discoverable by end users, but have decided that these repositories don't merit being included in the K-map.

Under certain conditions you might prefer to implement Discovery Server without a K-map, taking advantage of the spiders' thoroughness so that end users can find information simply by searching, but without taking the ramp-up time needed to create the full K-map. Such an effort would get you up and running with Discovery Server much more quickly.

The Discovery Server search-only solution enables indexing and searching across repository types and computes the document values that factor into the search results score. Because categorization wouldn't be enabled, the searching for expertise and affinity-generation features of the full Discovery Server wouldn't be available.

Instead of using the Startup navigator in the Discovery Control Center, use the Maintenance navigator:

1. Click the Servers & Services link under the System section in the Maintenance navigator.
2. Open the Server definition form, then enable all the Discovery Services except for the K-map Building and Affinity Processing services.
3. Under the Data section, click the Repositories link, then click the Add Repository button. You can ignore the K-map Settings link under the Data section because, for a Search-only implementation, there's no need to be concerned with K-map Creation.
4. On the Data Repository definition form, specify location parameters for each data repository plus any detail specific to the repository type. Be sure to complete the field map information appropriate for each

repository type because K-map Indexing relies on the field maps.

Data Repository [Save & Close] [Cancel] [Help]

Definition [Set schedule] [View status]

Title: Patents

Note: For all but e-mail repositories, it's best to reuse their native titles since these titles will be visible to K-map end users.

Use for:

☐ Both K-map categories and search results

☒ K-map search results only

Type: Notes Database

Server: Mfox/RvnTest

Database: PAT2000.nsf

5. Choose the option K-map search results only.
6. Schedule each repository for spidering.

Note: The Metrics section in the Maintenance navigator can be ignored. You can specify the links in the People section, except for the Affinities link, if you intend to enable end users to search for profiles. Just remember that affinities are dependent on categories existing in the K-map.

Specifying a subset of users for affinity generation

During the early phases of deployment, it's best to roll out affinities to your organization gradually by generating them only for a particular subset of users with whom you can test the accuracy of the affinities. Regardless of whether your affinity publishing policy is to auto-publish or to require end user approval, it's worth vetting the appropriateness and accuracy of the affinities against a targeted subset of users.

Discovery Server 2.0.1 introduces the new Notes.ini variable DS_Metrics_Affinity_Worker_Groupname to let you select a subset of users when generating affinities. For example, you can set:

`DS_Metrics_Affinity_Worker_Groupname=DSUserStage1,DSUserStage2,DSUserStage3`

where DSUserStage*n* includes each subset of users. As you grow your implementation, you add groups to the variable. The values entered can contain a comma-delimited list of groups as shown above or a single group name (to which you can add more users or groups as you increase the size of the deployment). The group or groups must already exist in the Domino Directory (and its equivalents) and can contain nested groups.

When assigning an affinity, the Affinity Processing service checks to see if the user is a member of the group that can receive affinities regardless of the affinity policy. If the group name is left blank, the Affinity Worker performs the actions dictated by the affinity policy to all users.

Deleting affinities for the purpose of starting over

There may be a need to regenerate affinities, for example, to respider a specific repository to improve the meta-data/authorship or to add another repository for consideration. Discovery Server 2.0.1 enables deleting affinities without having to reinstall the server, thereby losing all the work invested in building the K-map and does not require the respidering of all data.

Setting the Discovery Server's Notes.ini variable DS_Metrics_Affinity_Worker_Recompute=1 instead of 0 instructs the Affinity Processing service to perform a one time purge of any existing affinity state instead of proceeding with the current processing. During the affinity reset, the Affinity Processing service:

- Deletes all calculated affinity metrics, but leaves raw metrics intact.
- Deletes proposed and published affinities from all profile documents.
- Removes all pending affinity requests (proposals, approvals, rejections, removals) except for any profile edit requests from end users.
- Removes all users from the People Who Know About lists in K-map, although they might continue to display in these lists until the K-map cache clears. To clear the list, you can either wait for the information in the cache to expire (the default is 60 minutes after it was placed in the cache) or restart the primary Discovery Server and secondary servers where a K-map replica is enabled.

Note: After completing the reset, the DS_Metrics_Affinity_Worker_Recompute variable resets to zero.

Conclusion

This article briefly describes new features introduced in Discovery Server 2.0.1, the latest release of the Lotus knowledge management server. These features include accessibility-enabled K-map, better meta-data handling, search-only spidering and servers, and better control over affinity generation. We hope you found this article helpful—let us know about other Discover Server topics you'd like us to cover.

ABOUT THE AUTHOR

Jim Riel is a technical writer who has worked at IBM and Lotus for over six years.