Voyager Search Admin Guide

Version 1.9.3

Table of Contents

How Voyager Works 5

Data Discovery 5
Viewing and Searching the Index 5

Initial Setup and Configuration 6

Configuring Discovery Locations 6
Folder Locations 6
Web Services 6
ArcGIS Online or Server Account 7
Open Geospatial Consortium Catalog Services (CSW) 7
Indexing Data 8
Viewing the Index 9

Managing Voyager 10

Managing Discovery 11

Status 11
Locations 12
Directory 13
Web Services 13
AGS Indexing 13
CSW Indexing 14
WMS Indexing 14
Database Indexing 14
Managing Locations 17
Changing Location Display Name 17
Watch File System for changes 17
Schedule Location 17
Location Filters 17
Queue or Remove a Location 17
Location Discovery Settings 18
Exclude List 18
History 18
Schedule 18
Formats 18
Discovery Configuration – Default Settings 19
Formats 20
Settings 20

Thumbnails 21
Discovery Configuration – Processing Settings 21
Metadata Extraction 22
HTTP Client 26
Field Mapping 27

Managing the Index 27

Query Settings 27
Query Synonyms 28
Dynamic Labeling 29
Manage Solr 29
Federated Catalog Search 29
Backup and Restore 31
Clear Index 31

Managing Settings 32

Voyager URL 32
Managing Voyager's Appearance 33
Localization 34
Site Title 34
Site Color 34
Banner 34
Site Footer 34
Disable Eternal Links 34
Adding a Custom Logo 35
Editing the Appearance Settings File (appearance.dex) 35
ArcGIS 36
Mapping 36
Logging 37
License 38

Managing Extensions 39

Managing Security 39

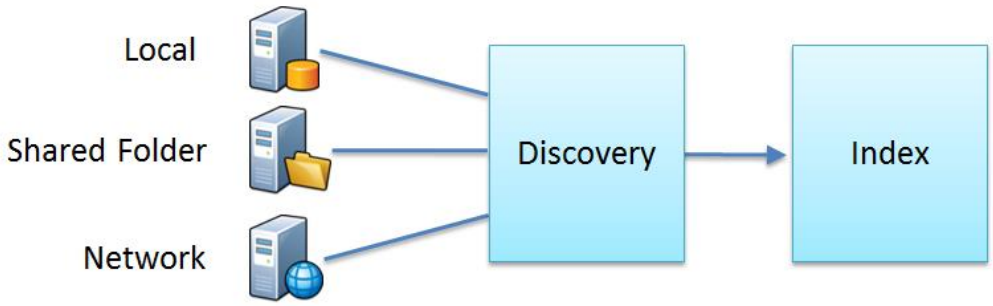
Authentication 40
Internal Authentication 40
Adding a Group 41
Windows Authentication 42
Windows Authentication Groups 43
Single Sign-On 43
LDAP Authentication 44
LDAP Authentication Groups 44

Sessions 45
Permissions 45
Managing Users 46
Managing Access 47
Query-Based Access Restrictions 47
Location-Based Access Restrictions 48
Managing System Settings 50
Special Admin Options 51
Remove Results from the Index 51
Pending Extraction 52
Show Raw Query 52
Data Inventory and Assessment 53
Finding Duplicates 53
Missing data links 53
Most commonly used data 53
Distributed Searching 53
Distributed Data Discovery 53
Index Replication (Load Balancing) 53
Index Sharding 54
Web Application Clustering (Load-Balancing) 54
Hardware and Software Requirements 54
Hardware 54
Operating Systems 54
Web Browser 54
ESRI Data Indexing Requirements 54
Virtual Machines 55
Additional Software 55
Additional Considerations 55
Access to data 56
Adding a Custom Task to Voyager 57
Creating the info file 57
Creating the Python Script 58

How Voyager Works

Voyager crawls content in specified discovery locations, creates an index for the results and provides comprehensive search capabilities within that index. Searches are highly configurable, as are the various display modes. You can save searches as well as process the results for further analysis.

Voyager does not move or copy data; it only creates an index for data that it discovers.



Data Discovery

Voyager ingests content in a series of steps in the *Indexing Pipeline*. The Indexing Pipeline makes two passes over the data as it builds the index.

- The first pass identifies all available data in the specified locations(s), and typically takes very little time to complete. It reads different data formats using *connectors* specific to each MIME type. The connector framework is extensible and new data types can easily be added.
- In the second pass, Voyager gathers detailed information using *extractors* for different file types. During this phase, it uses metadata where it is present, but does not require it to complete the indexing. It also reads any other available information for each record, such as geospatial location or any keywords associated with a map. Voyager then assembles all of this information in the index and creates thumbnails for each record. Depending on the size of the data set, this pass can take considerably longer to complete.

Viewing and Searching the Index

Once Voyager has completed both passes, it displays the results in the Summary window. From here you can view the results in different layouts, search the results, save searches and process the data further.

Initial Setup and Configuration

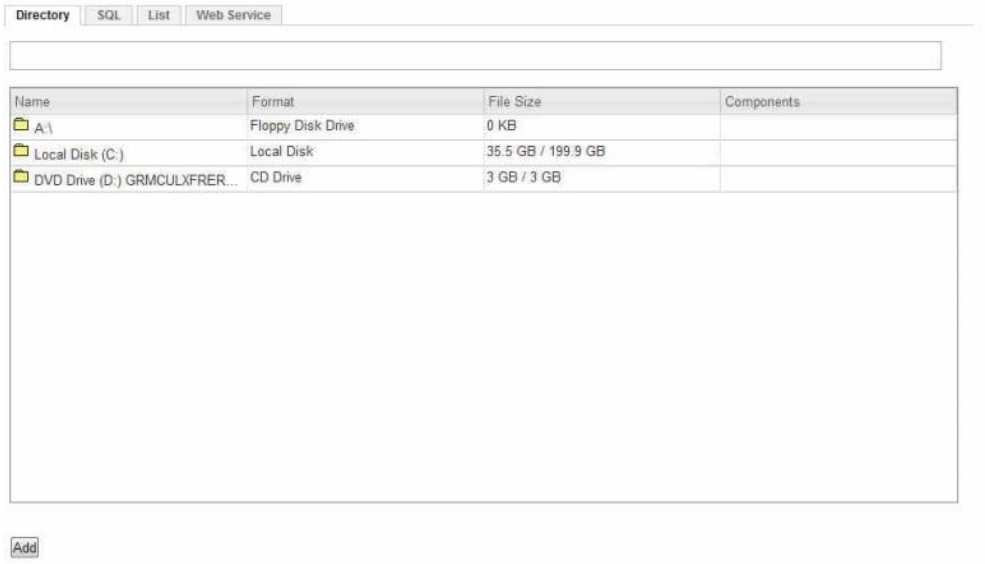
Voyager's configuration wizard walks the user through the initial process.

Configuring Discovery Locations

Voyager's first task is to discover the data in your collections, which can be stored locally or on a server.

Folder Locations

This step requires specifying a folder location that Voyager inspects for data and documents.



By clicking a drive location, you can browse your file system for directories to index. You can also type the relative path in the search bar at the top. Once a folder is selected, clicking the **Add Location** button includes it in the list of locations to be indexed.

Web Services

In addition to adding file-based locations, you can also index Web Services. To add a Web Service location, select the **Web Service** tab.

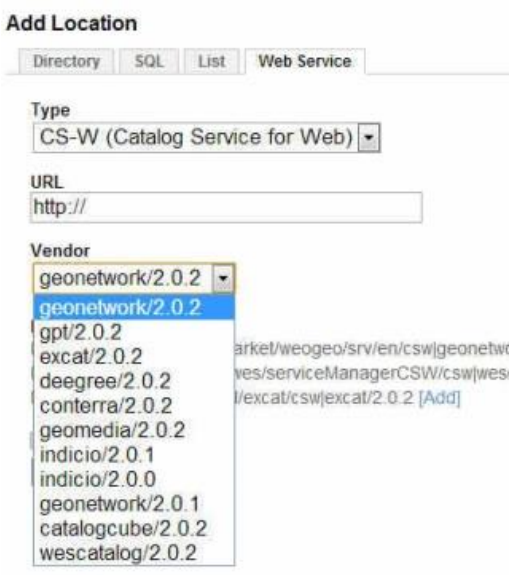
For a Web Mapping Service, add the URL, check **Validate** and click **Add**.

ArcGIS Online or Server Account

For an ArcGIS Online or Server Account, type in the URL and enter authentication details.

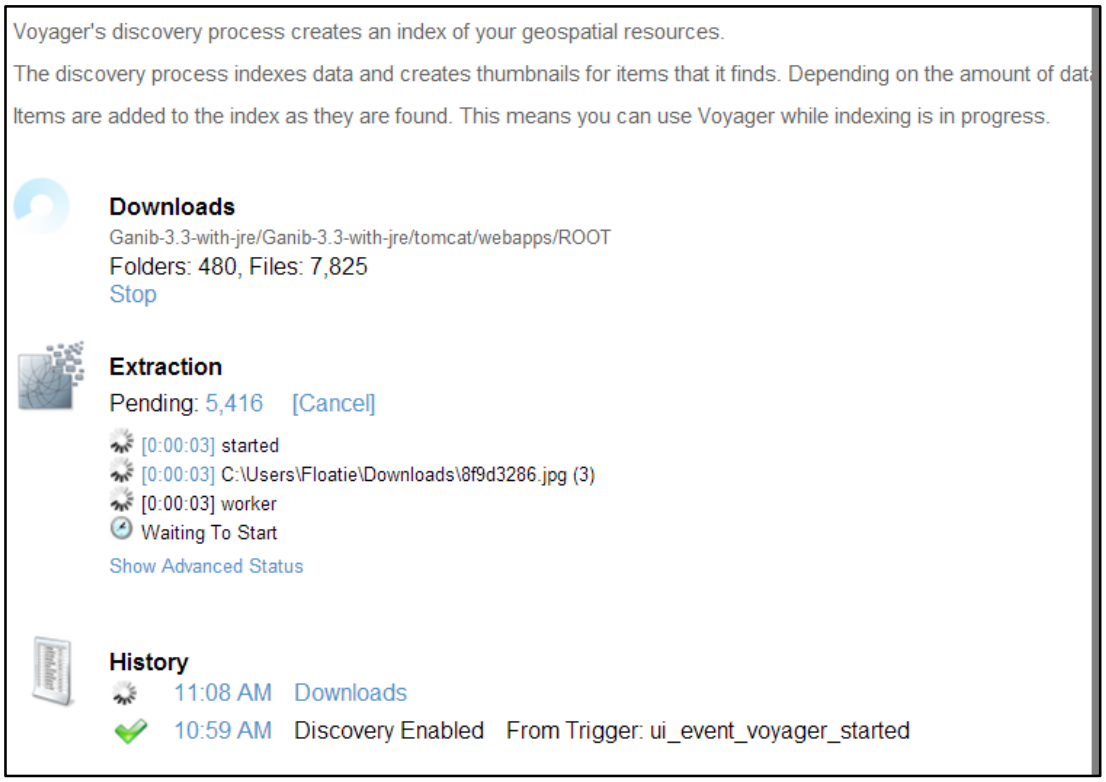
Open Geospatial Consortium Catalog Services (CSW)

CSW compliant servers can be indexed by adding the server URLs to Voyager's list of Discovery Locations. To add a service enter the URL, select the CSW server software used to publish the service and then press Add. Note that a few example services are provided.



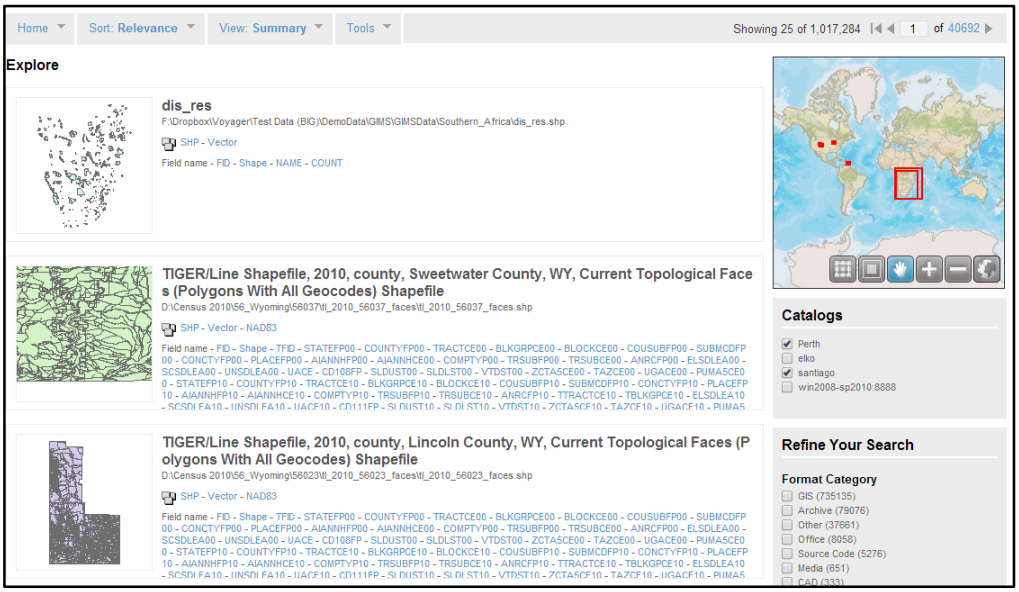
Indexing Data

After you have added discovery locations, Voyager will begin to crawl the data and construct an index.



Viewing the Index

When indexing is in process, you can view the results in the Summary window. Note that it can take a few minutes for items to appear.



Managing Voyager

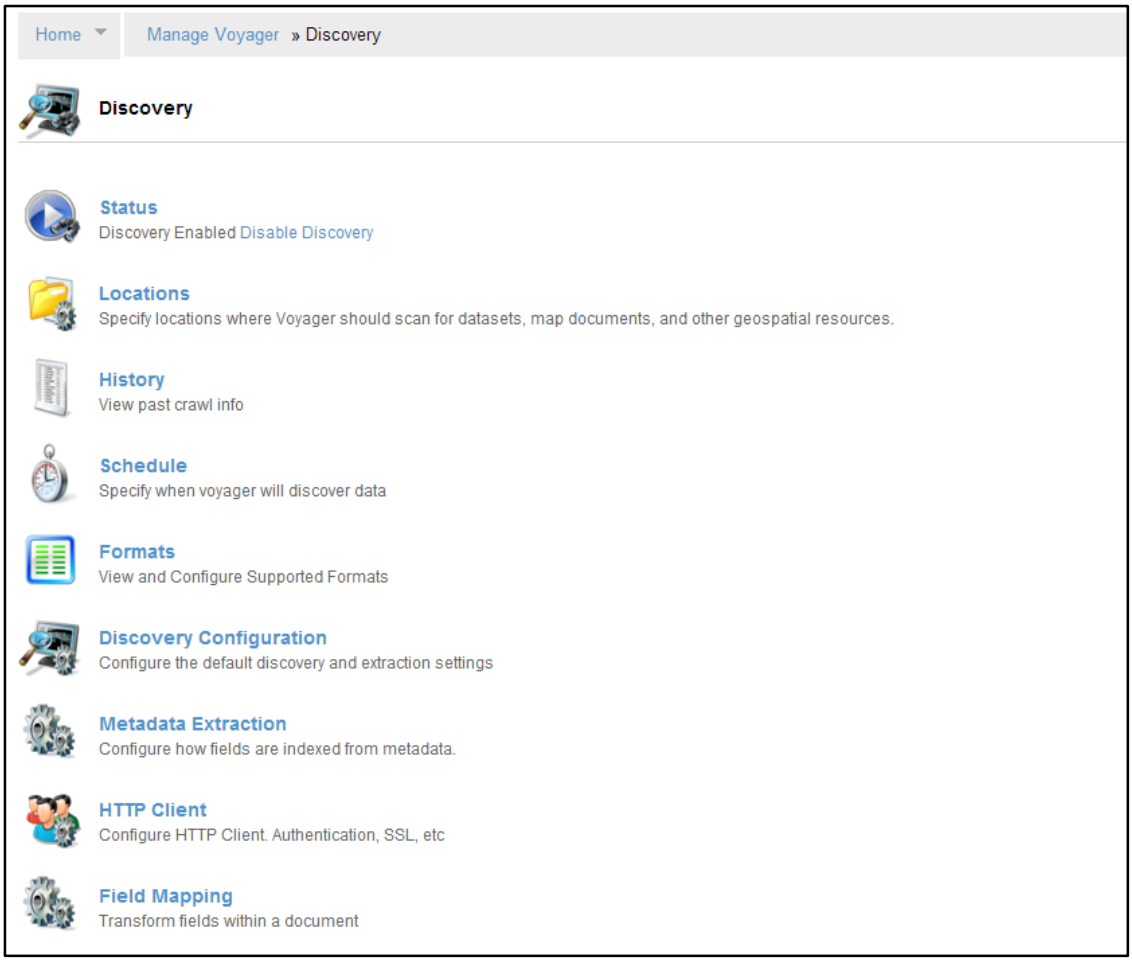
This Administrator Guide describes how to manage and configure Voyager, including information on discovery, the index, general appearance and mapping settings, extensions and security settings.

These features allow the user to adjust and configure the Discovery, Settings, Security, and System portions of Voyager as well as check for software updates. Navigate to these features by clicking **Manage** in the upper right-hand corner.

Note: If the option to manage Voyager is not displayed, you may have to first login. Voyager can be configured to require administrative access to manage these settings. The default administrative account name is 'admin' and the password is 'admin'.

Managing Discovery

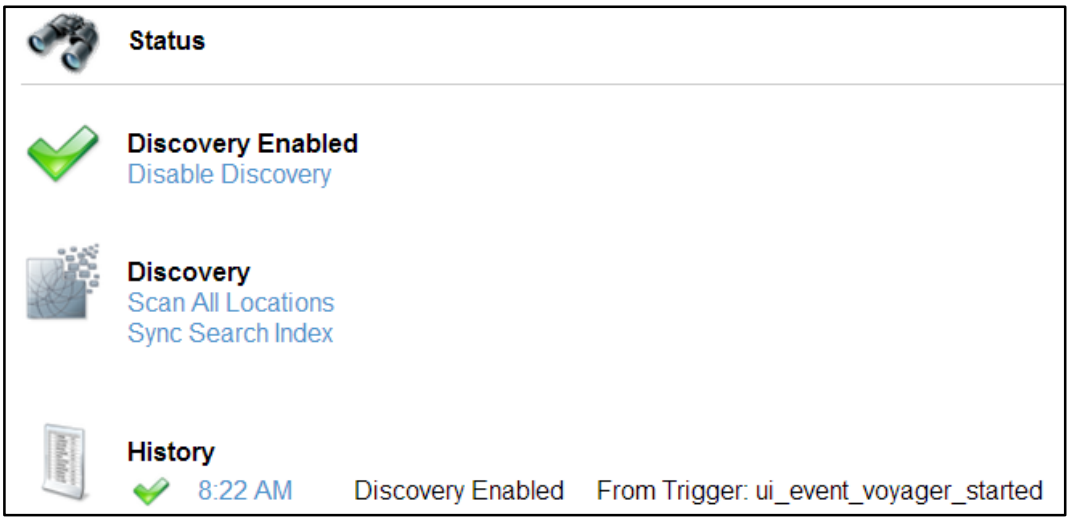
To manage all of the aspects of the discovery process, go to **Manage Voyager > Discovery**.



Status

[**Manage Voyager > Discovery > Status**]

The Status page indicates if the discovery process is enabled, and shows the number of pending tasks in the indexing and thumbnail queues.

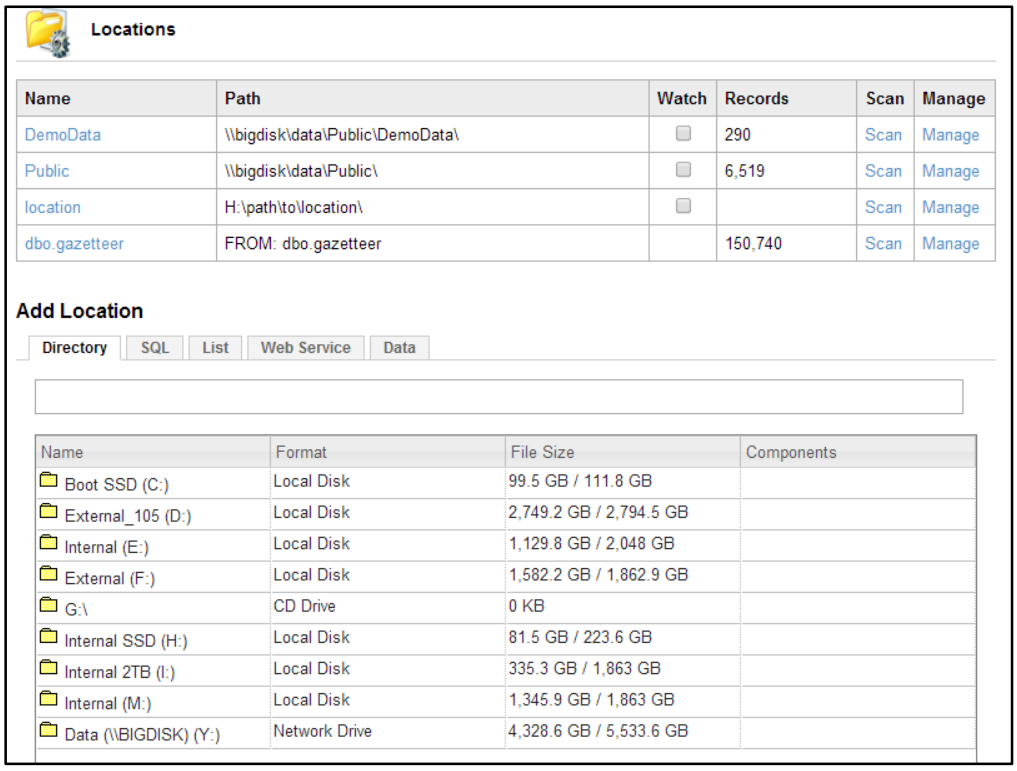


- To start or stop indexing, selecting **Enable** or **Disable Discovery**
- To rescan all of the configured discovery locations click **Scan All Locations**
- During indexing, Voyager caches results so the items may not immediately viewable in the index. **Sync Search Index** pushes all items in the discovery cache into the viewable index.
- **History** displays a list of discovery events, for example when discovery is enabled or disabled, indexing status and any errors that occur.

Locations

[**Manage Voyager > Discovery > Locations**]

The Locations page is used to specify where Voyager should scan for datasets, map documents and other resources. Depending on your license level, locations can either be local, on a network, or web-based.



Location information is displayed and contains the name, path, indexing schedule, count of items discovered at this location, as well as an option to queue the location and manage the discovery settings for this specific location.

Directory

You can add a directory or folder can be added as a location. For example you could add D:\data or server\share as valid locations. If you are new to Voyager, it is best to start with a discovery location that contains only a few datasets. As your experience grows, you can add larger amounts of data.

Web Services

In addition to indexing items on disk, Voyager also indexes web services. To add a web service to the list of discovery locations, click the **Web Service** tab on the **Locations** page. Enter the URL to a web service and choose whether it is WMS (Web Map Service), AGS (ArcGIS Service), or CSW (Web Catalog Service). Choosing **Validate** will immediately verify whether the server exists at the specified URL and will report whether Voyager can successfully extract information from the web service.

AGS Indexing

ArcGIS Server web services can be added to the index from the Locations page. Enter the URL to the services directory on the ArcGIS Server, select **AGS** from the dropdown and choose whether to immediately validate the URL. This will index all supported services on this ArcGIS Server.

It is also possible to enter a URL to a specific directory or service on the server. The image below shows three valid ArcGIS Server URLs. The first is to an ArcGIS Server (ArcGIS Online), the second is to a folder within the server (ArcGIS Online - Elevation Directory), and the third is to a specific map service (World_Street_Map).

CSW Indexing

Web Catalog Services can be added to the index from the Locations page. Enter the URL to the CSW, select **CSW** from the dropdown and choose whether to immediately validate the URL.

WMS Indexing

Web Map Services can be added to the index from the Locations page. Enter the URL to the WMS, select **WMS** from the dropdown and choose whether to immediately validate the URL.

To more efficiently add WMS locations to the index, .wms files can be leveraged. URL(s) can be added to a .wms file, then the directory containing the .wms file can be added as a Location. Voyager will open the .wms file and index all URLs contained within it.

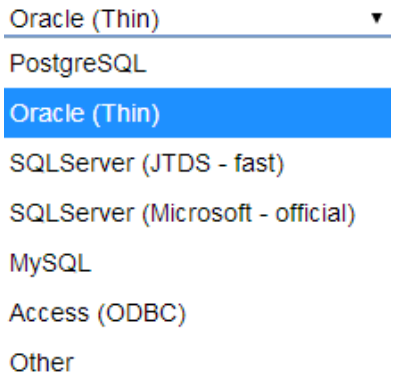
Database Indexing

[Manage Voyager > Discovery > Locations > SQL]

Voyager 1.9.x adds the ability to index and search the rows of a database without requiring any knowledge of the database schema. Voyager provides some demo data for you to explore, however we recommend that you consult your database administrator before you implement this feature.

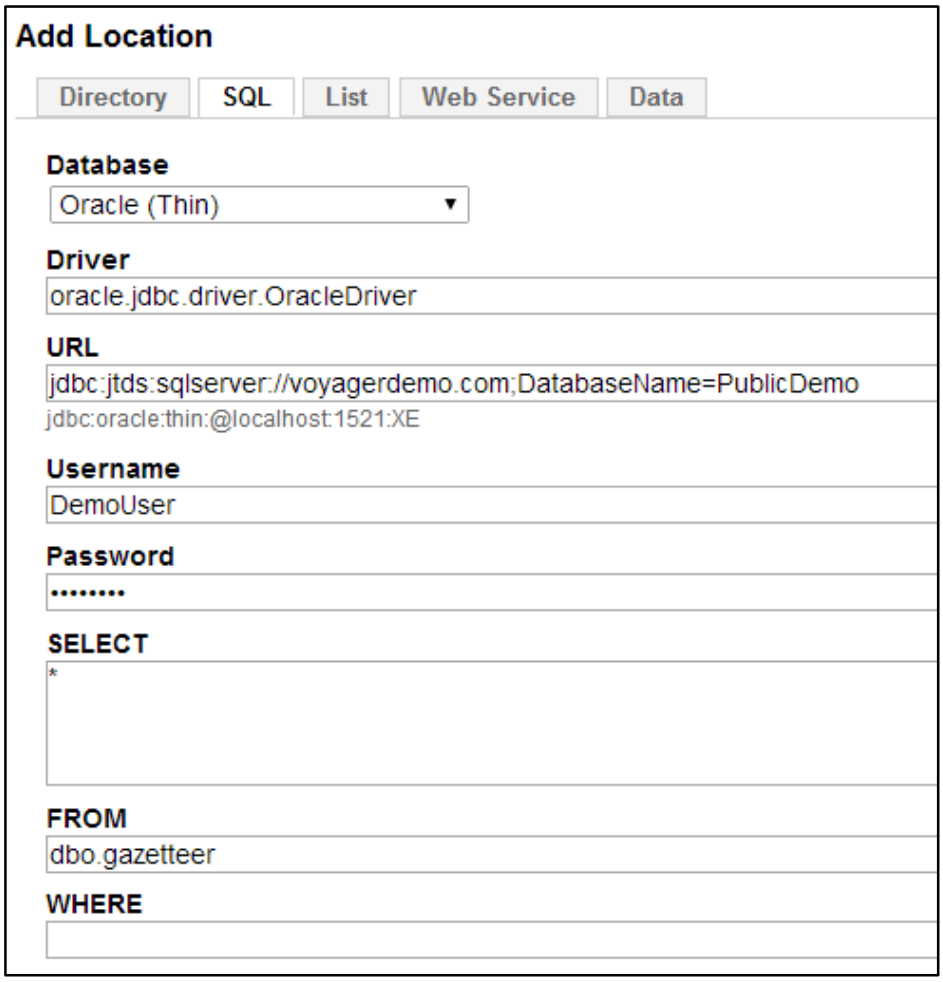
To add a database to the indexing process:

1. Go to **Manage Voyager > Discovery > Locations**
2. Click the **SQL** tab
3. Select one of the options in the Database drop-down list, for example Oracle (Thin), as shown below:



The values in the **Add Location** dialog depend on the type of Database you choose from the list. For the demo databases, Voyager fills in these fields for you. If you are adding your own database, your administrator can configure these values to match your environment.

In the following example, **Oracle (Thin)** is the selected database type.

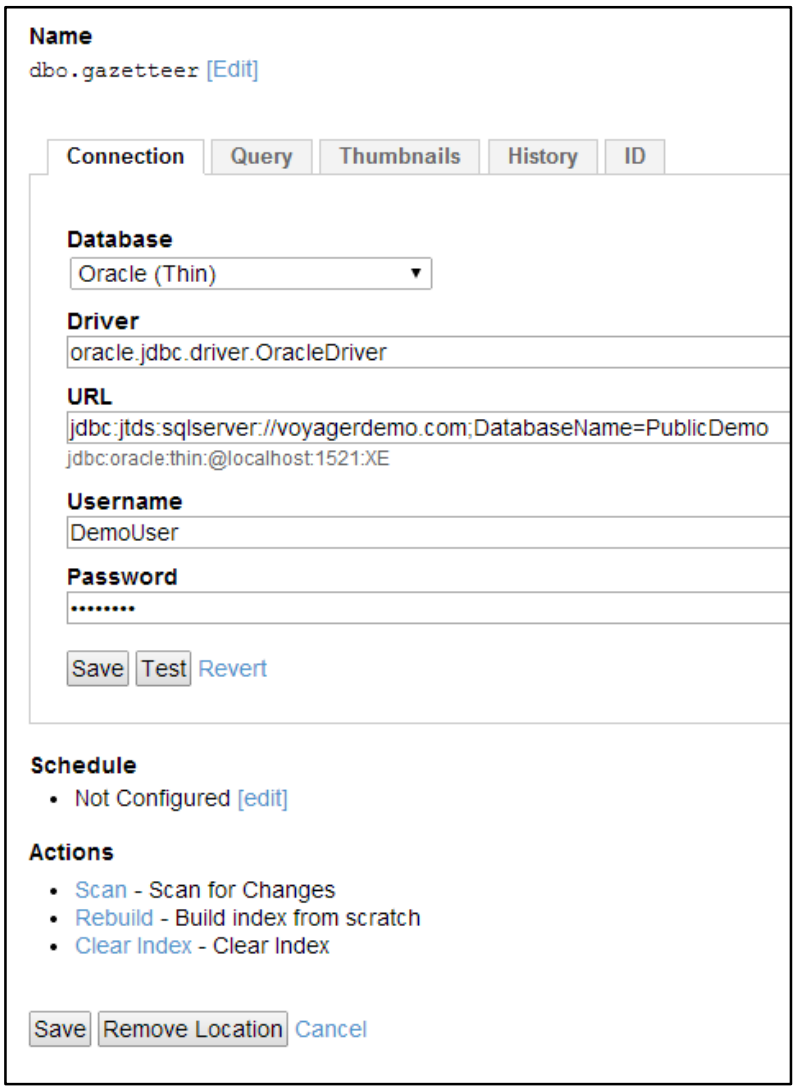


1. **Driver** is the jdbc driver that Voyager will use to connect to the database

2. **URL** is the jdbc connection path to the database
3. Enter a **Username** and **Password** if required
4. **SELECT** specifies the data to be extracted. The default value is all data (indicated by ***)
5. **FROM** identifies the database owner
6. **WHERE** specifies subsets of the data to be indexed
7. When you are done, click **Add**

This brings up a **Connection** dialog where you can modify, test and schedule the discovery settings.

1. To add the database connection to the locations list, click **Save**
2. To begin indexing, click **Scan**



You can see more information about the scan **at Configure > Voyager > Discovery > Status.**

Managing Locations

The discovery settings for each location can be configured using the location's **Manage** link. A dialog is displayed showing the various options to configure. The settings configured here will apply only to this location, not to the overall index.

Changing Location Display Name

The display name of the location can be changed by clicking on the name, which brings up an editable field where the name can be modified.

Watch File System for changes

You can choose to **Watch** a location for any changes made to the directory on the file system. The Voyager index will be updated immediately if a file in the watched location changes.

Watching a location will only work on files that have a 1-to-1 mapping to files on the system. For example, it works well for file-based Rasters, Shapefiles, Office documents, MXDs, Layer Files, etc., but will not currently work for content stored in SDE, VPF, or anything else that has a different internal structure than what is displayed as a file on disk.

Schedule Location

The discovery process for a specific Location can be scheduled to run automatically. Choose from an existing schedule task and add it to the location.

Location Filters

Configuring filters allows you to choose which types of data you want Voyager to search for. Filters are used to include or exclude files or folders that are found within this location.

For example, you could choose to add only MXDs within this location to the index. To do this, choose **Add Filter** under **Include**, decide whether it should be case sensitive, and enter *.mxd for the Pattern. The click the **Add Filter** button and the filter will be saved for this location. To exclude items from the indexing process, follow the same steps under the **Exclude** section.

When excluding items, be sure to create any exclusion filters before indexing the location. If a location has already been indexed, any items that should be excluded that have been previously indexed will remain in the index.

Queue or Remove a Location

The location can be added to the **Discovery Queue** by clicking **Queue Location**. It can be completely removed from the list of locations by clicking **Remove Location**.

Location Discovery Settings

Each location can override the default **Data Discovery Settings**.

Exclude List

During the discovery process, you may want to exclude certain system folders, file types, files with certain prefixes, etc. from the index - regardless of which discovery location the item is in. The **Exclude List** is a global list of things Voyager will skip when indexing.

For example, you may want to ensure that items within the C:\Windows folder are never indexed by Voyager. To do this, click **Exclude List**, then **Add**. A dialog is displayed allowing you to configure what to exclude.

To exclude the C:\Windows folder, select **Folder** from the dropdown menu and choose C:\WINDOWS from the list. Click **Exclude**, and the folder you selected will be added to the Exclude List.

To remove the location from the **Exclude List**, click the **[x]** and the folder will no longer be excluded from the discovery process.

In addition to excluding system folders, you can exclude items based on wildcards, or regular expressions (regex). For example, if you want to exclude everything with the prefix test_ from the index, select the **Wildcard** option and enter test_* as the **Pattern**.

History

[**Manage Voyager > Discovery > History**]

History displays a list of discovery events, for example when Discovery is enabled or disabled, indexing status and any errors that occur.

Schedule

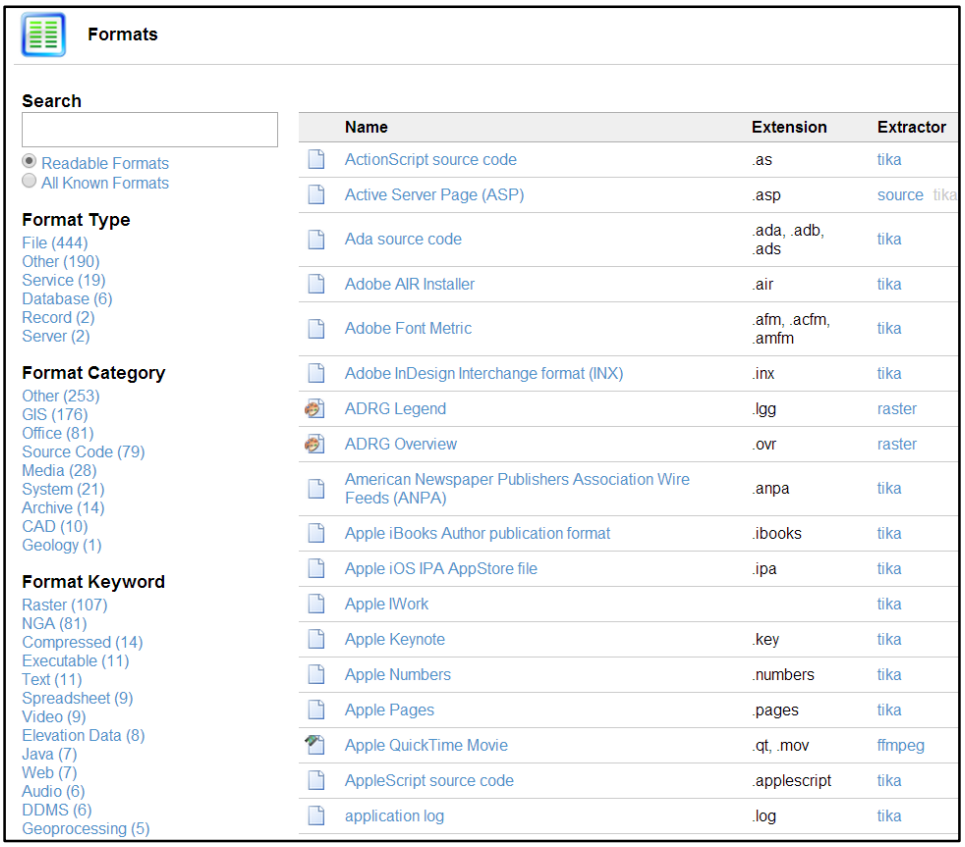
[**Manage Voyager > Discovery > Schedule**]

Use the **Schedule** feature to schedule data discovery. On the Scheduling page, you can add triggers, which configure the discovery process to run based on time of day, periodically, chronologically or on the occurrence of specific events.

Formats

[**Manage Voyager > Discovery > Formats**]

This option shows the list of supported file formats (MIME types). The example below shows only the top of the file list — the complete list is very much longer.



- Click the **Extractor** name to see all of the file types associated with that extractor
- Click the **Name** to see a detailed description of that file type and the extractors that can read it
- You can also use the **Search** box at the top to find format information
- Click any of the filters at the left to refine the view of files. These include
  - Format Type
  - Format Category
  - Format Keyword
  - Format Company
  - Extractor
  - Format Application
  - Product

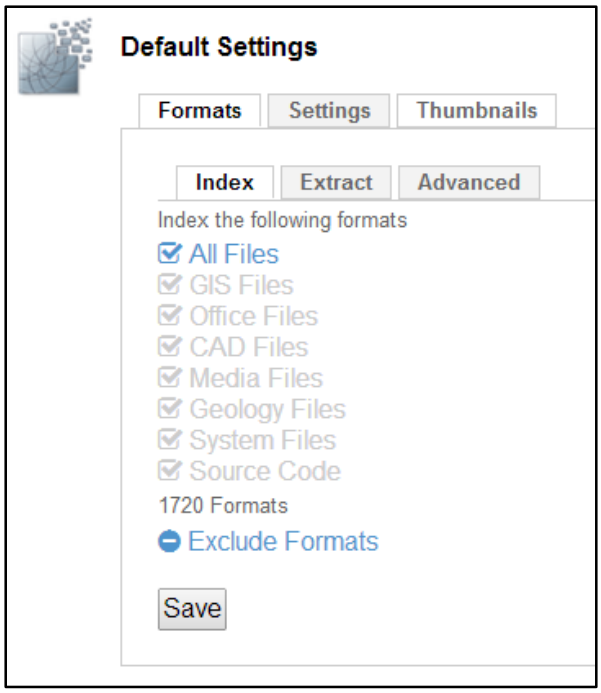
Discovery Configuration – Default Settings

[**Manage Voyager > Discovery > Discovery Configuration**]

Use this option to configure discovery and extraction settings as well as limits for cpu and RAM usage. In the top part of this page, you can configure default settings, e.g. include or exclude particular file formats.

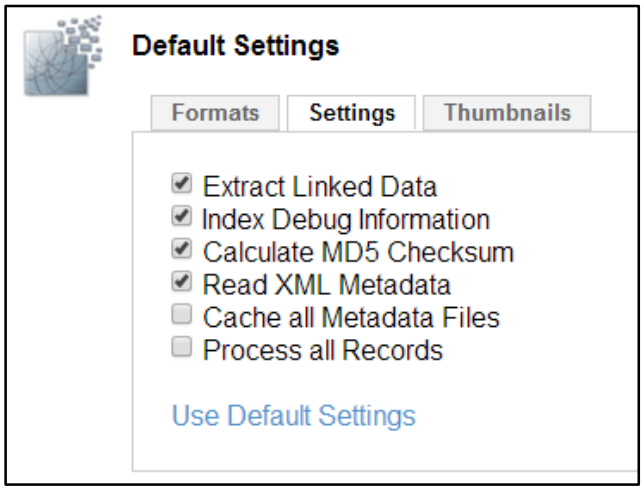
Formats

Select Formats to manage indexing and extraction options for different file formats, as well as relevant MIME type options.



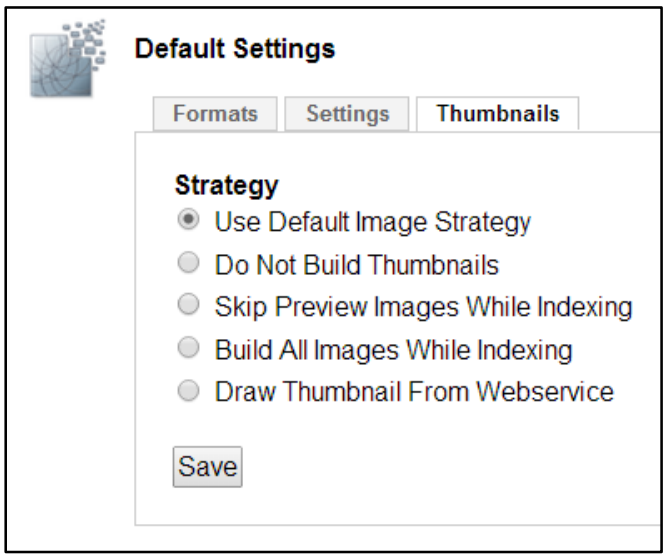
Settings

Select **Settings** to configure some indexing and debug settings.



Thumbnails

Use the Thumbnails tab to configure how Voyager generates thumbnails.



Use Default Image Strategy

With this setting, Voyager builds thumbnails on demand. This means that Voyager does not build thumbnails for index records that never appear in a search. Keep in mind that if the system is very busy, some results may not have thumbnails until the indexing process catches up.

Do Not Build Thumbnails

This allows the fastest indexing, since no thumbnails are generated.

Skip Preview Images While Indexing

When it indexes a file and extracts information, Voyager creates a thumbnail and a preview for that file and stores them in the Meta folder in your Voyager directory. When you select this option, Voyager does not generate any previews.

Build All Images While Indexing

This option builds thumbnails for all results and is generally the slowest option.

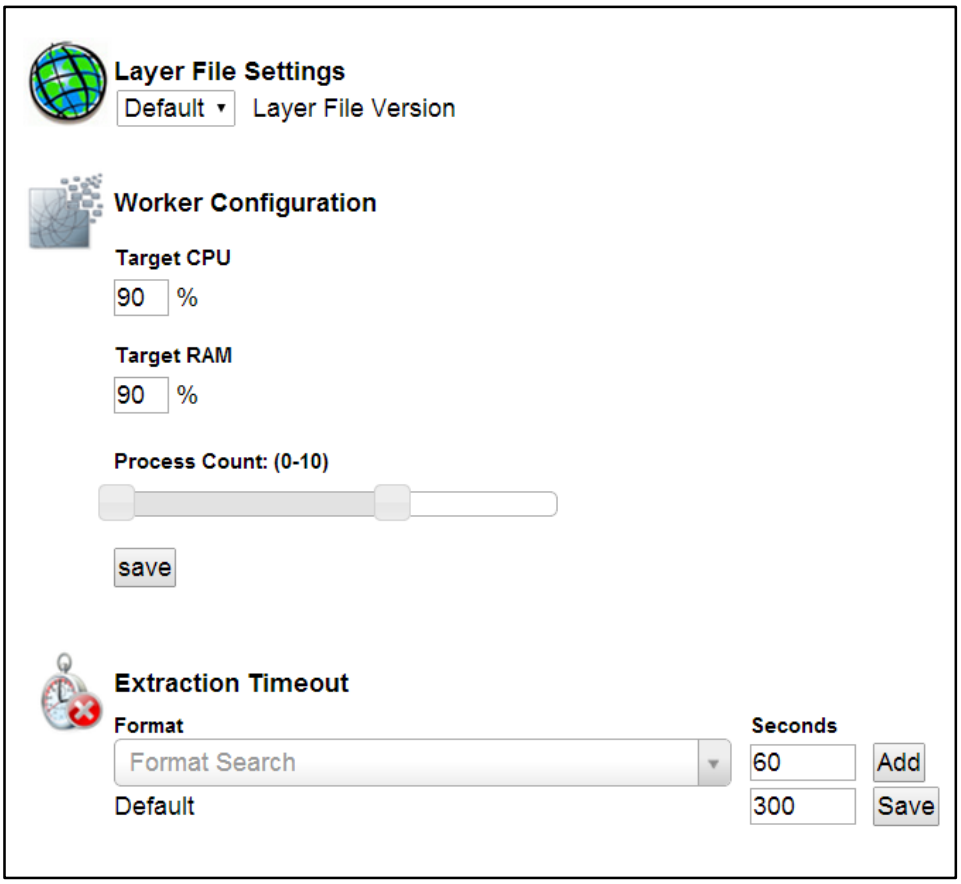
Draw Thumbnail From Web Service

Use this option to have the indexing progress select a thumbnail image appropriate for the file type.

Discovery Configuration – Processing Settings

[**Manage Voyager > Discovery > Discovery Configuration**]

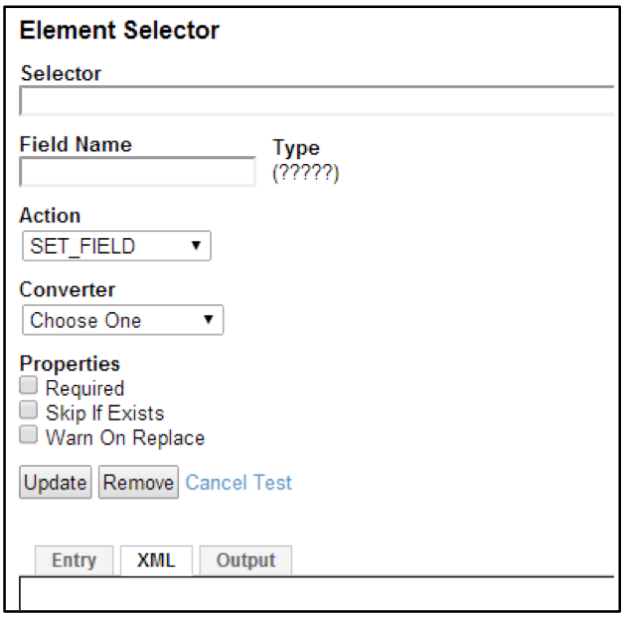
On the bottom half of the page, you can configure limits for processor and RAM usage as well as set the timeout values for data extraction.



Metadata Extraction

[**Manage Voyager > Discovery > Metadata Extraction**]

When Voyager is indexing content from different content stores, it is sometimes useful to map the varying schemas into a common set of values. Voyager handles this requirement through **Metadata Extraction. Metadata Extraction** pulls metadata from standard XML documents using XPath queries. Voyager's out-of-the-box setting support many standard metadata specification, but it also allows you to enter your own XPath queries to metadata elements and map them to searchable field names within Voyager's index.



To map the fields, configure these parameters:

Selector

This specifies XPath query to a specific metadata record element to be selected.

Field Name

This is the target field in Voyager that gets mapped to the specified metadata output.

Type

This refers to the data type of the field name. For example, if field name is set to name, data type automatically gets set to text.

Action

You can select from five different functions:

- **Set Field**— Assigns value to the specified field
- **Append Field**—Adds to/modifies pre-existing field
- **Set Geo**—Sets a geographic bounding box based on coordinates specified in the metadata
- **Expand Geo**—Expands geographic bounding box from previously set coordinates
- **Add Link**—Helps point field to a URL

Converter

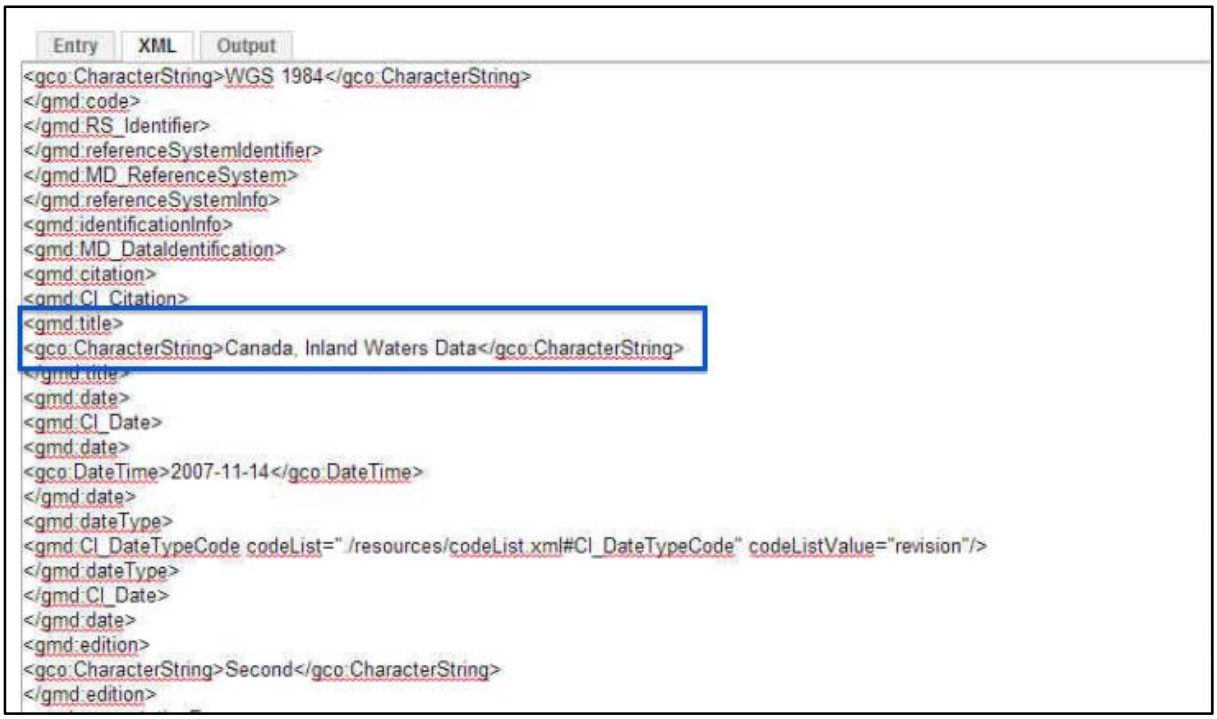
Converter settings are optional. If you do not specify a particular setting, Voyager assigns an appropriate default converter to the field.

- **Bbox** — Converts bounding box values contained in the XML document
- **Gml_Geometry**— Converts geometric (line, circle etc) coordinates from the XML document
- **Date** — If the Date field is represented as a string value in the XML document, this converts it into a standard date format
- **String256** — Finds a String within the element with a maximum length of 256 characters
- **String512** — Finds a String within the element with a maximum length of 512 characters
- **StringValue** — Finds a String (of any length) within the element Properties
- **Required** — Checking this box validates the field being extracted from the XML document.
- **Skip if Exists** — If a field has been previously added, checking this box ensures that a duplicate field does not get added to your list.
- **Warn On Replace** — If a destination field already exists, checking this field flags the newly set field value.

Using the XML Box

The XML box allows you to enter in an XML document to test your XPath queries to paired elements. Click the XML tab and paste the contents of a valid XML document here. Click **Save** to save the XML contents.

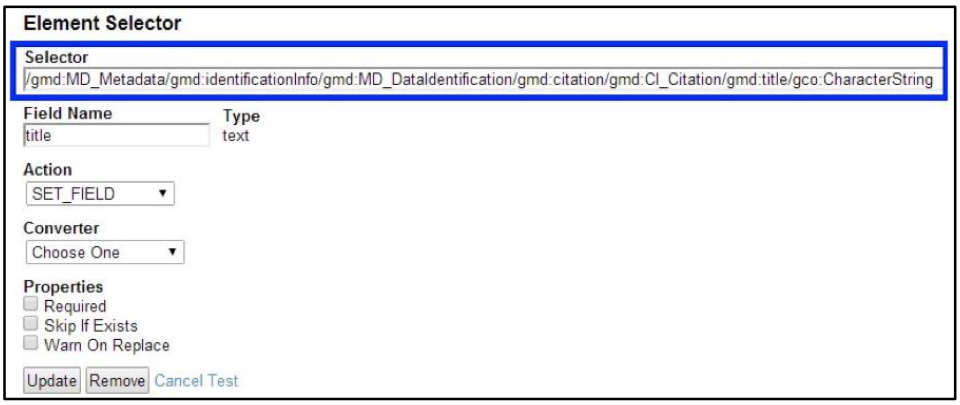
In this case, the element we want extracted from the XML tab is **Title**.



Specify values for **Selector**, **Field Name** and **Action**.

Since we want to extract the field **Title**, we copy the XPath Query from the XML document in the Selector box.

/gmd:MD_Metadata/gmd:identificationInfo/gmd:MD_DataIdentification/gmd:citation/gmd:CI_Citation/gmd:title/gco:CharacterString.



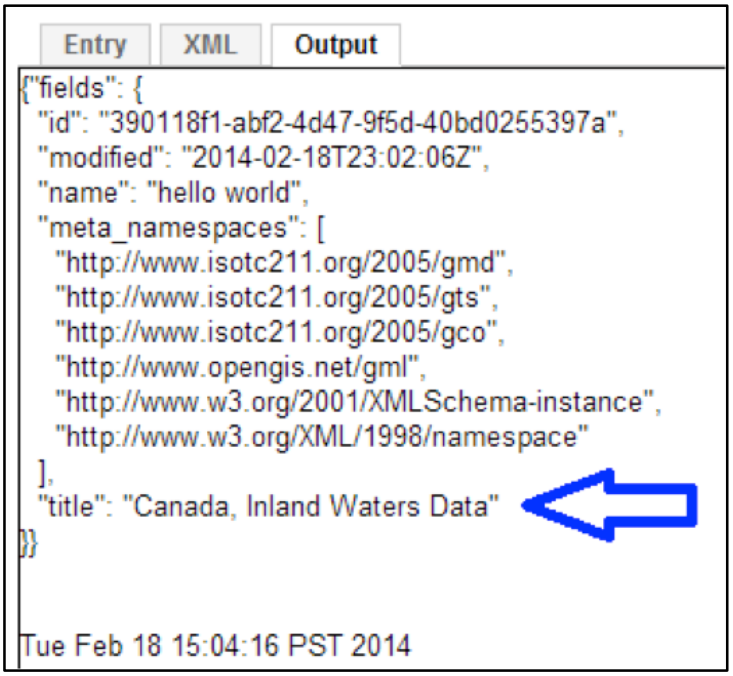
Specify the corresponding **Field Name** to be mapped to the queried element. Voyager automatically detects the **Type** for the **Field Name**.

For example, here the Field Name is **Title**, whose field type is **Text**.

Note: when selecting a field name you'll need to either select an existing field name or you can also enter a custom field name as long as it uses a prefix meta_, id_.

Click **Test**. The extractor searches the XML document for the queried metadata element, and retrieves the value for the field **Title**. The results are presented in the **Output** tab.

In this specific example, Canada, Inland Waters Data, which is the value for the Title query, is retrieved from the XML tab and displayed in the Output tab. Users can use this output result to search for XML documents through Voyager's search UI.



- Click **Save** to add the XPath query to the list
- Click the **Edit** link to make changes to an existing Selector
- Use the up or down Arrows to change the order of a Selector
- Select the [X] to delete an existing Selector

HTTP Client

[**Manage Voyager > Discovery > HTTP Client**]

Use this option to configure the HTTP Client. Authentication, SSL, etc

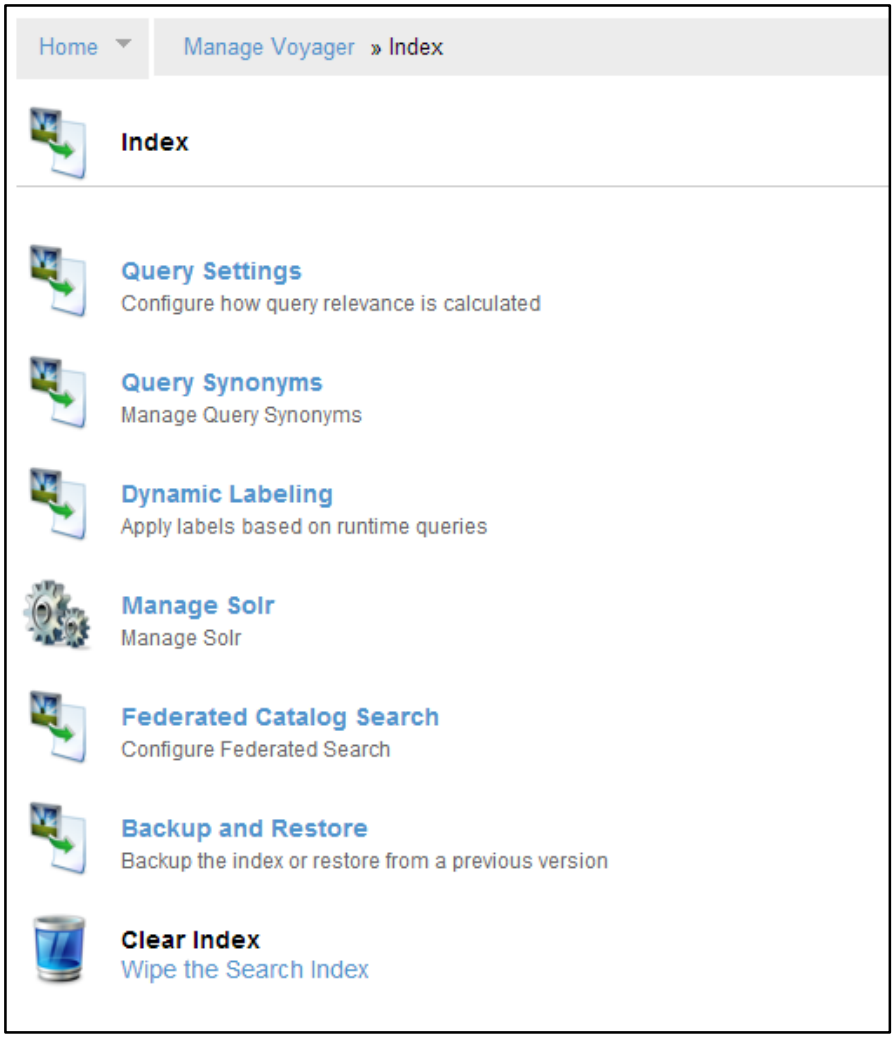
Field Mapping

[**Manage Voyager > Discovery > Field Mapping**]

This is where you can map and transform fields within a document.

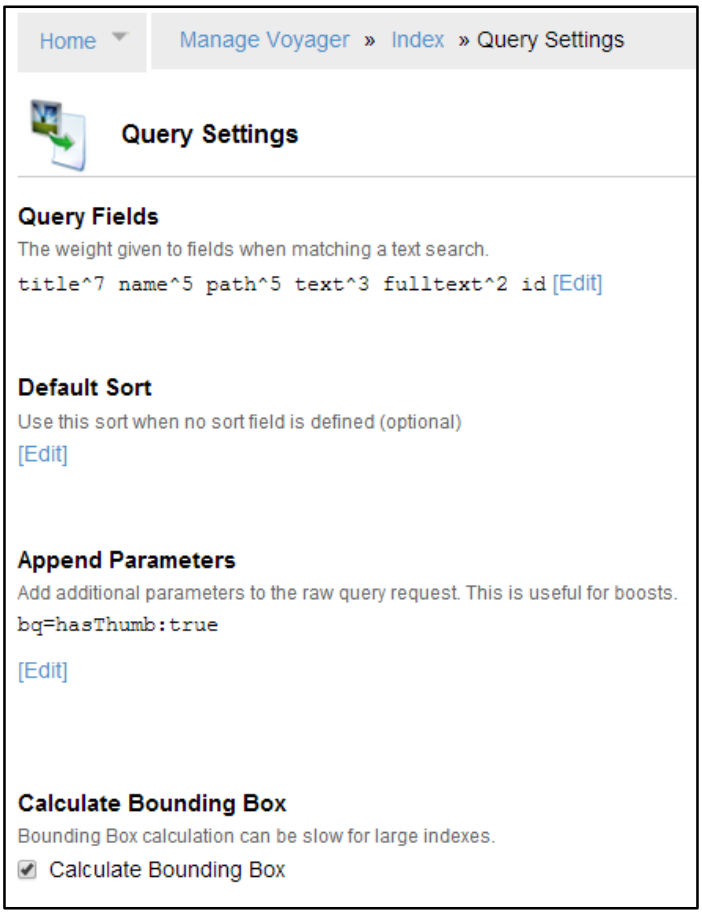
Managing the Index

Use this section to manage features of the Indexing process. Go to **Manage Voyager > Index**.



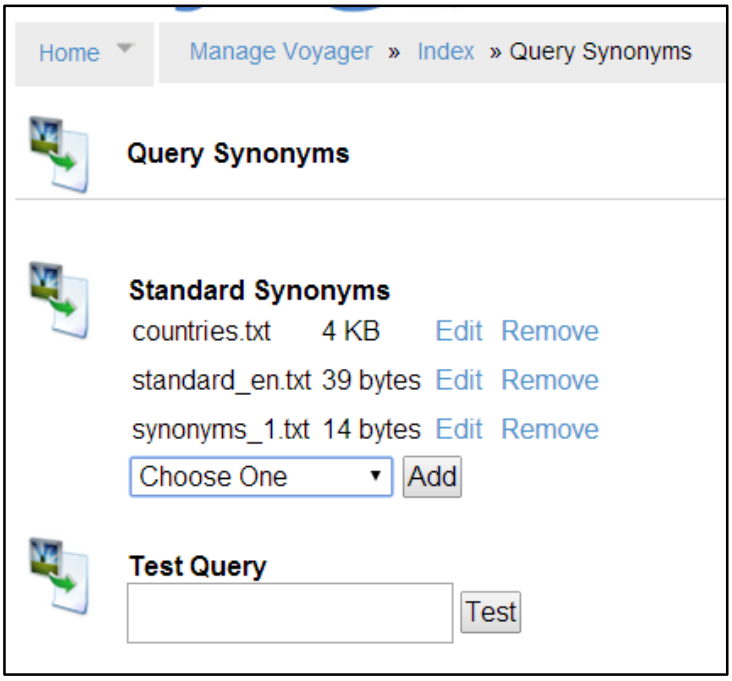
Query Settings

[Manage Voyager > Index > Query Settings]



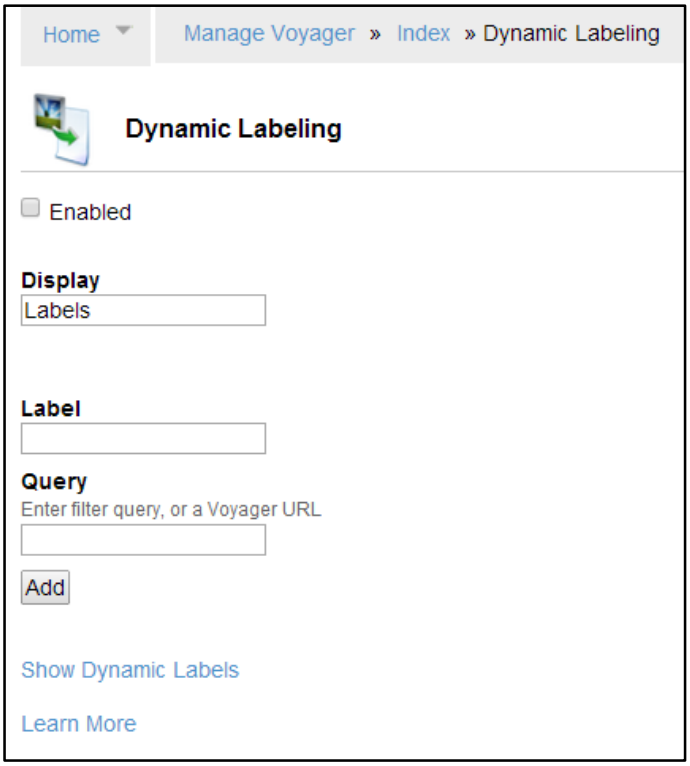
Query Synonyms

[Manage Voyager > Index > Query Synonyms]



Dynamic Labeling

[Manage Voyager > Index > Dynamic Labeling]



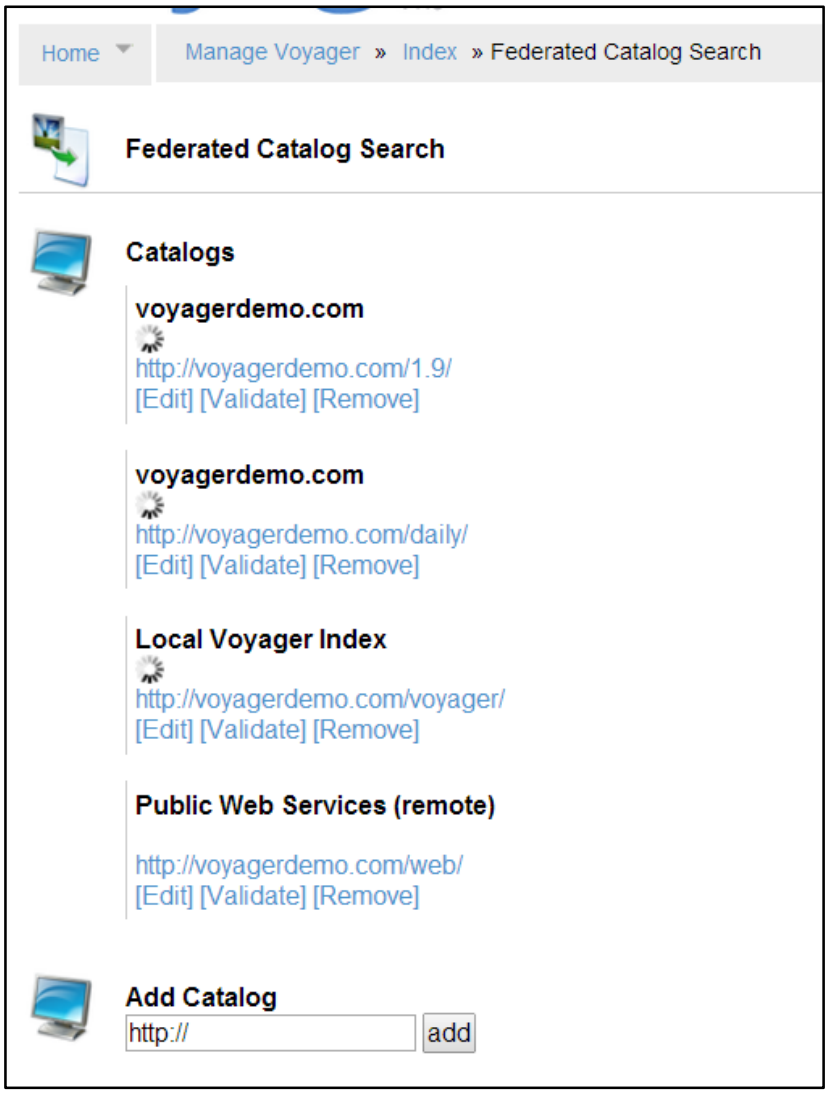
Manage Solr

[Manage Voyager > Index > Manage Solr]

Federated Catalog Search

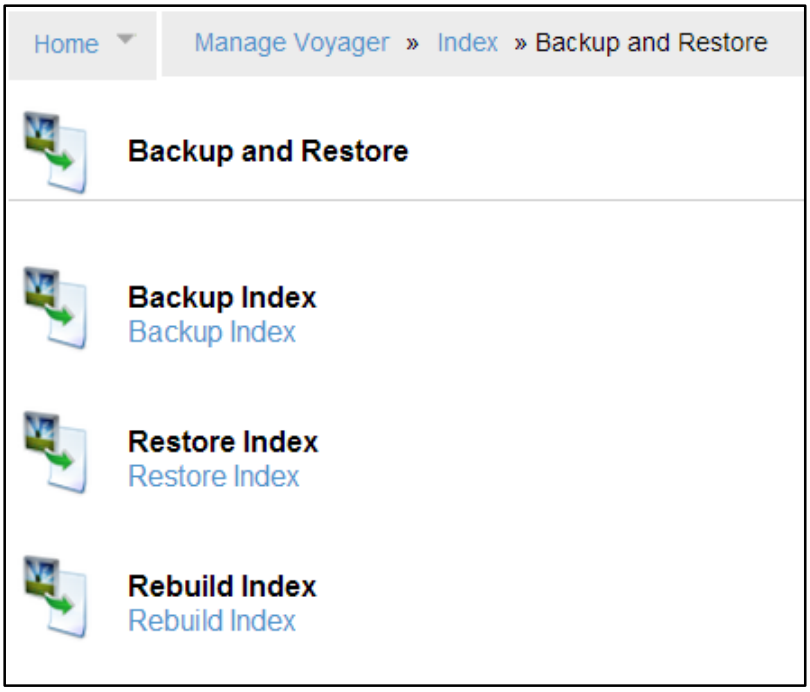
[Manage Voyager > Index > Federated Catalog Search]

Users can simultaneously search multiple Voyager instances. Federated Search helps organizations link multiple Voyager installations and indices to provide a single comprehensive view of their entire spatial data infrastructure.


Home
Manage Voyager » Index » Federated Catalog Search
Federated Catalog Search
Catalogs
voyagerdemo.com
http://voyagerdemo.com/1.9/
[Edit] [Validate] [Remove]
voyagerdemo.com
http://voyagerdemo.com/daily/
[Edit] [Validate] [Remove]
Local Voyager Index
http://voyagerdemo.com/voyager/
[Edit] [Validate] [Remove]
Public Web Services (remote)
http://voyagerdemo.com/web/
[Edit] [Validate] [Remove]
Add Catalog
http://
add

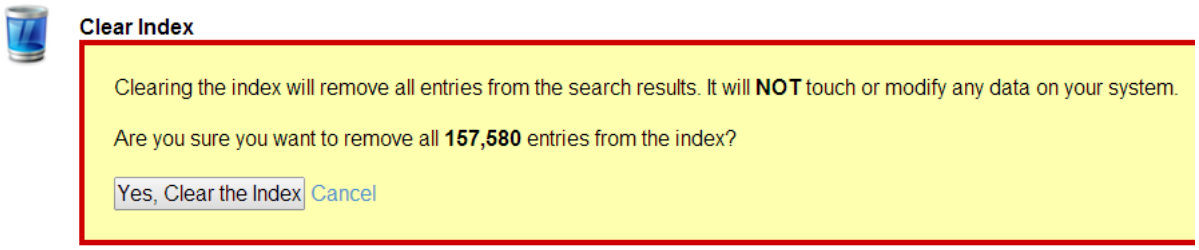
Backup and Restore

[Manage Voyager > Index > Backup and Restore]



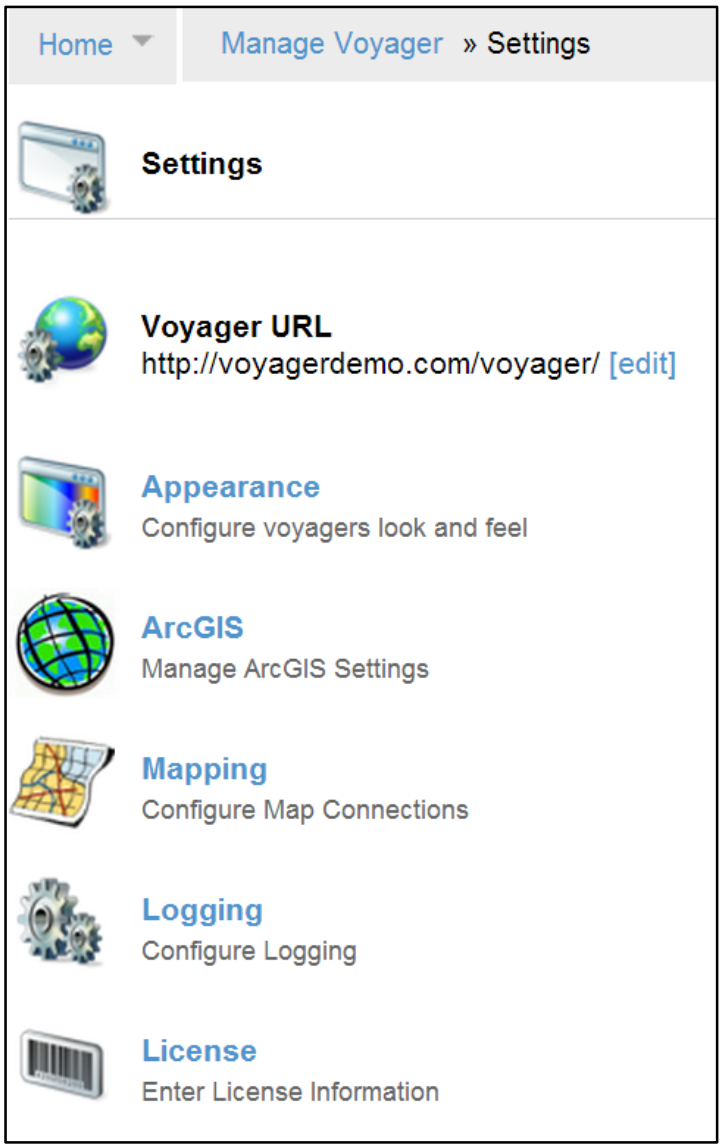
Clear Index

[Manage Voyager > Index > Clear Index]



Managing Settings

This section displays system and version information and lets you configure system logging as well other attributes. To view Settings information, go to **Manage Voyager > Settings**.



Voyager URL

[Manage Voyager > Settings > Voyager URL]

When you deploy Voyager as a server, you must set the Voyager URL so that web-based users can navigate to Voyager

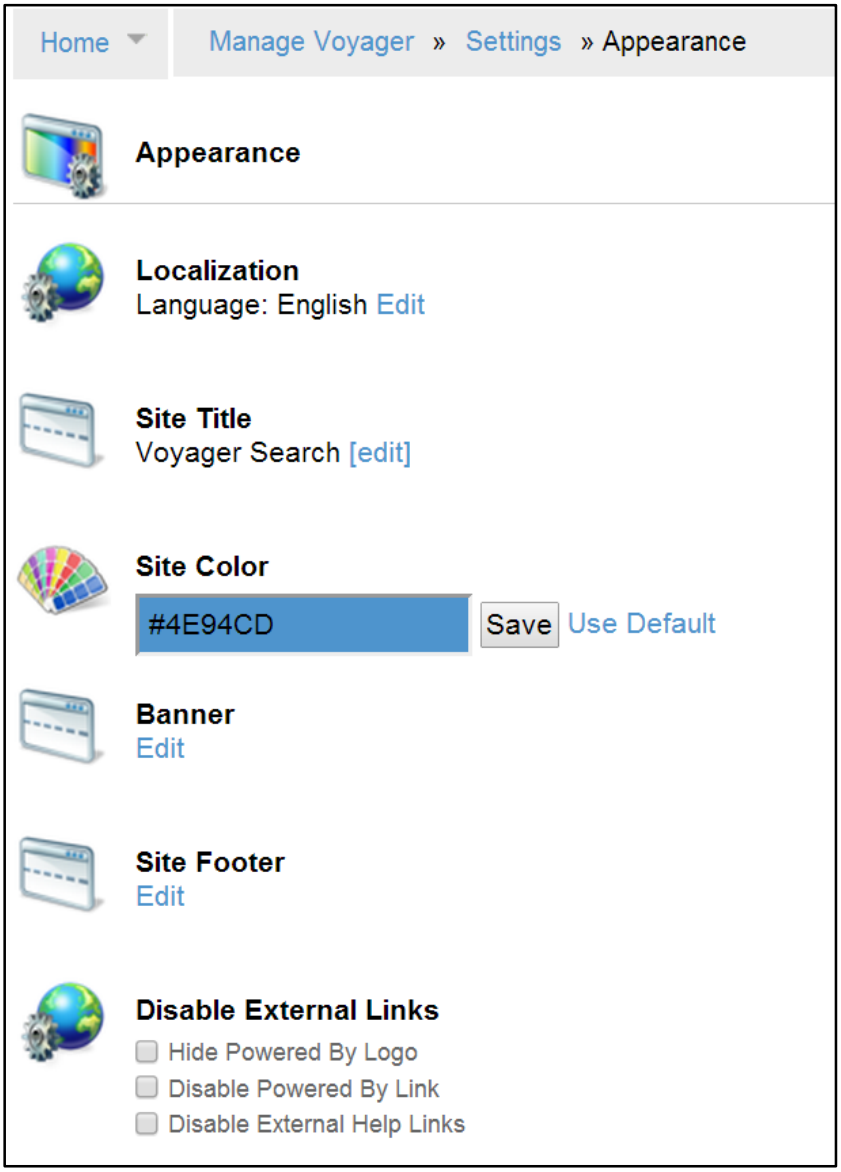
Click **Edit** to change the URL setting and enter a valid host machine name or web address.

This is required to support:

- Connecting via the Voyager ArcMap Dockable Window
- CSW Server
- Remote catalogs accessed from Federated Search
- RSS Feeds

Managing Voyager's Appearance

[Manage Voyager > Settings > Appearance]



Localization

To change the language settings, click **Edit** and choose from the drop-down menu.

Site Title

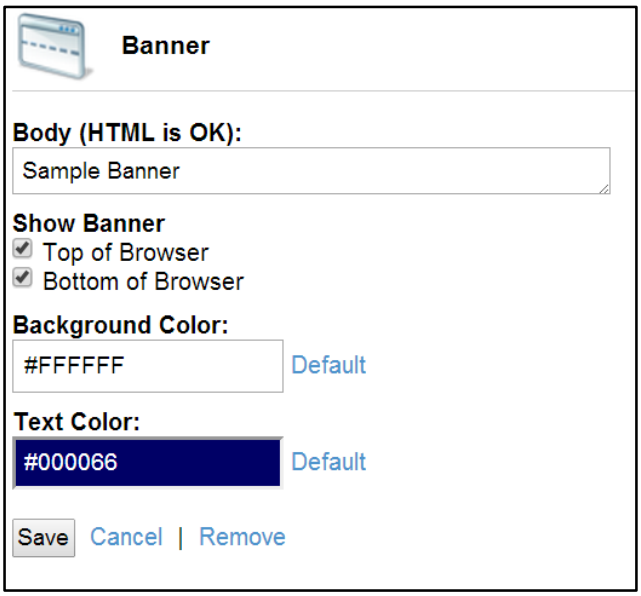
To change the site title, click **Edit** and enter a new value. Click **Save** when you are done.

Site Color

To change the site color, enter a color hex value directly or use the color picker to choose a new color. Click **Save** when you are done.

Banner

Click Edit to change the banner text and appearance.



Click **Save** when you are done.

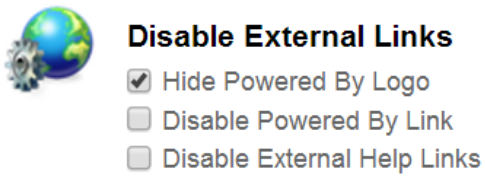
Site Footer

Enter the text you'd like to appear in the footer of each page. Click **Save** when you are done.

NOTE: If you have also added a banner to the bottom of the page, it will obscure the footer text. To view the footer text, uncheck **Bottom of Browser** in the banner settings.

Disable Eternal Links

You can hide the **Powered By: Voyager** logo at the bottom of each page or you can simply disable its link to www.voyagersearch.com.



Check **Disable External Help Links** to prevent Voyager from showing links to online help topics. When this is checked, clicking **Help** displays only the Voyager version.

Adding a Custom Logo

You can replace the Voyager logo with a custom logo of your own design. To add your logo to Voyager:

1. Save your logo as a .png file and rename it to **header.png**
2. Copy **header.png** to **${data.dir}/public**
3. Restart Voyager
4. To revert to the default logo, delete **${data.dir}/public** and restart Voyager

Editing the Appearance Settings File (appearance.dex)

Changes you make to appearance in the Voyager UI are stored in the file **${data.dir}/config/appearance.dex**, for example:

```
{
  "title": "NEW TITLE",
  "color": "#CD0829",
  "footer": "New Footer",
  "disableExternalHelpLinks": true,
  "banner": {
    "showHead": true,
    "showFoot": false,
    "body": "New Banner",
    "color": "#3D44FF",
    "background": "#FFFFFF"
  }
}
```

You can also edit this file directly. Your changes will appear when you restart Voyager. If you are deploying multiple Voyager instances and you'd like them all to look the same, you can create a standard **appearance.dex** file and copy it to the **${data.dir}/config** directory for each instance.

ArcGIS

[**Manage Voyager > Settings > ArcGIS]**

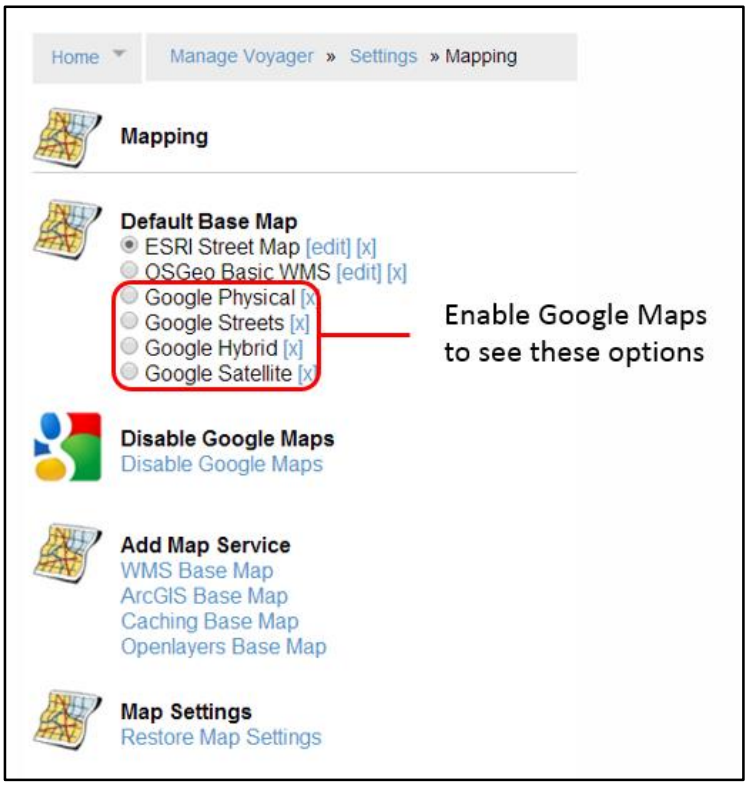
You can manage ArcGIS license settings here.

Mapping

[**Manage Voyager > Settings > Mapping]**

In the Mapping dialog, you can

- **Select a Default Base Map**
  ESRI Street Map, OSGeo Basic WMS or Google Maps (if enabled)
- **Enable Google Maps**
  Select this option to add items to the Default Base Map list.
- **Add Map Service**
  - **WMS Base Map**
  - **ArcGIS Base Map**
  - **Caching Base Map**
    Select this to add a cached map that can be accessed when offline
  - **Openlayers Base Map**
- **Restore Map Settings** to the default

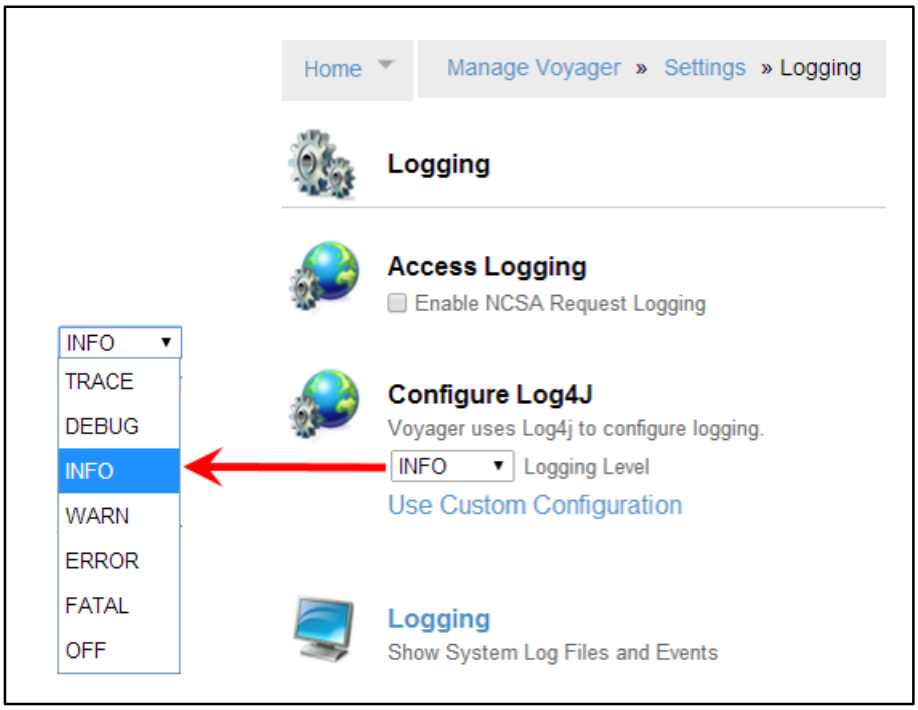


Logging

[Manage Voyager > Settings > Logging]

Manage logging settings and view current logs.

- **Access Logging**
  Enable NCSA Request Logging to log incoming requests.
- **Configure Log4J**
  This is where you select a logging level, from OFF (no logging) to TRACE (complete logging). You can create your own custom logging configuration.
- **Logging**
  View the current logs.



The logging levels are:

Level	Description
OFF	No logging
FATAL	Logs severe errors that cause premature termination
ERROR	Logs runtime errors or unexpected conditions
WARN	Logs situations that are undesirable or unexpected
INFO	Logs interesting runtime events (startup/shutdown)
DEBUG	Logs detailed information on the flow through the system.
TRACE	Logs the most detailed information

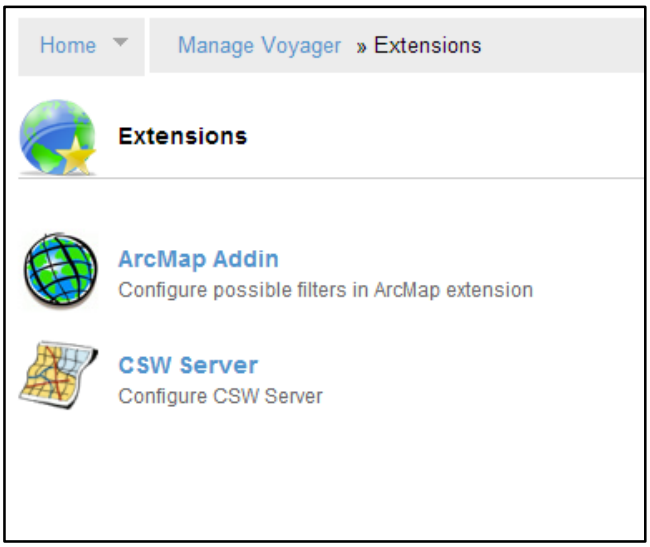
License

[Manage Voyager > Settings > License]

View your current license or enter a new license.

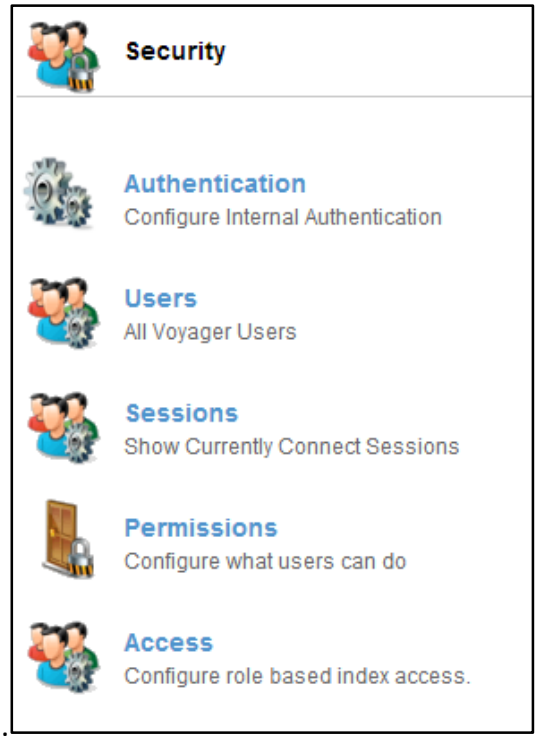
Managing Extensions

Use this section to manage Extensions to Voyager as well as the ArcMap Add in and CSW Server. Go to **Manage Voyager > Extensions**.



Managing Security

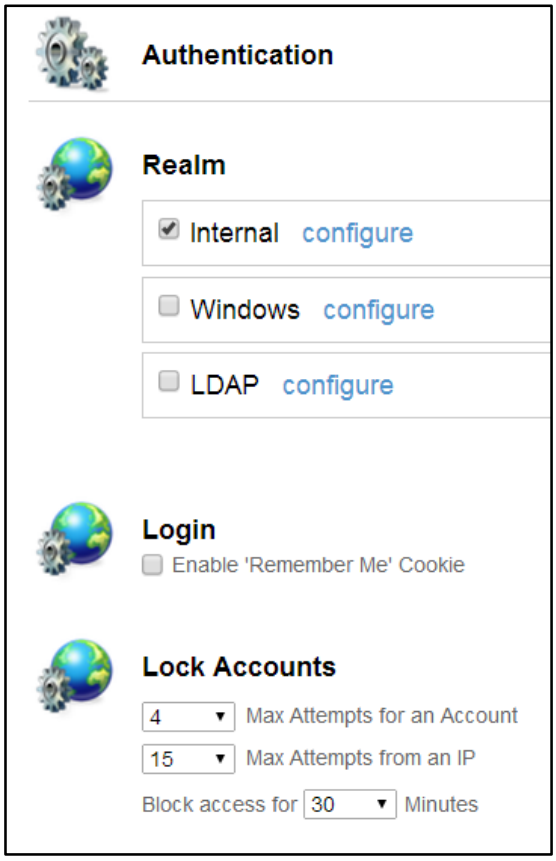
Here you can manage all aspects of authentication, permissions and access as well as display Voyager users and current sessions. Go to **Manage Voyager > Security**.



Authentication

[**Manage Voyager > Security > Authentication**]

Use the authentication page to choose an authentication model. You can choose from **Internal**, **Windows** or **LDAP** authentication.



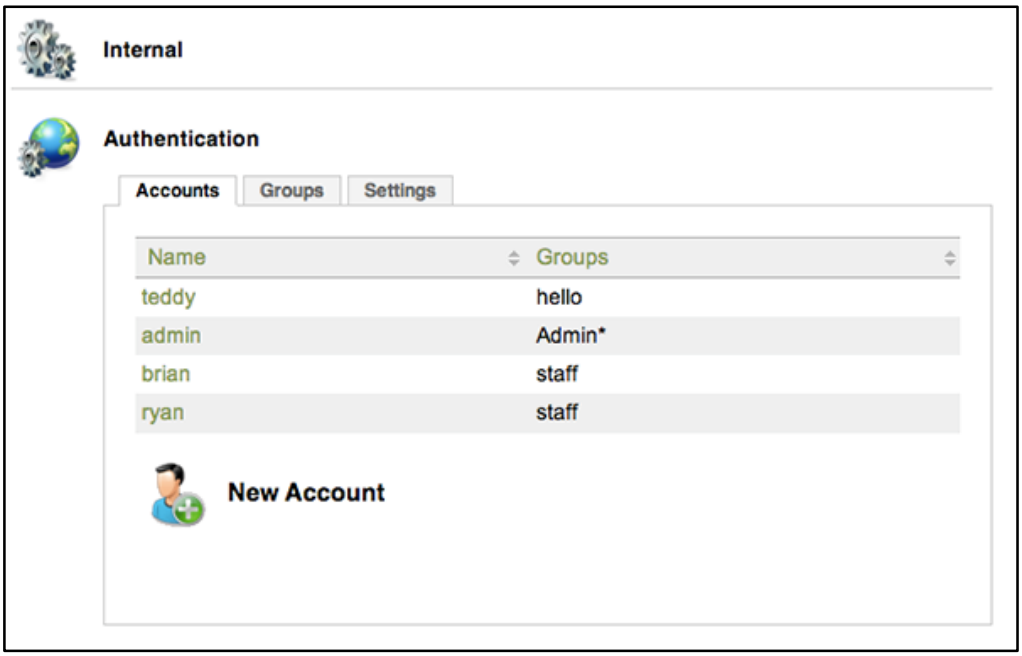
Remember-me authentication allows Voyager to remember the identity of a person between sessions.

Lock Accounts specifies when and how long to lock an account after incorrect login attempts.

Internal Authentication

[**Manage Voyager > Security > Authentication > Internal**]

Voyager's default authentication model uses configurable users and groups to control how people access and use Voyager.

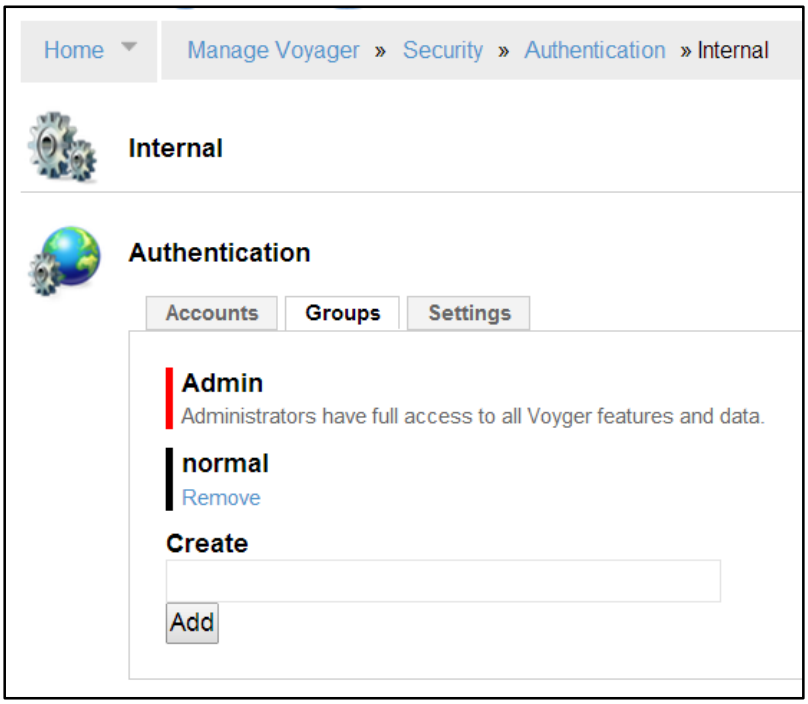


Click **Settings** to configure password length and complexity.

Adding a Group

To add a group using internal authentication:

1. Go to **Manage Voyager > Security > Authentication**
2. Select **Internal** and click **Configure**
3. Select the **Groups** tab
4. Enter a new group name in the **Create** field and click **Add**.



Windows Authentication

[**Manage Voyager > Security > Authentication > Security > Windows**]

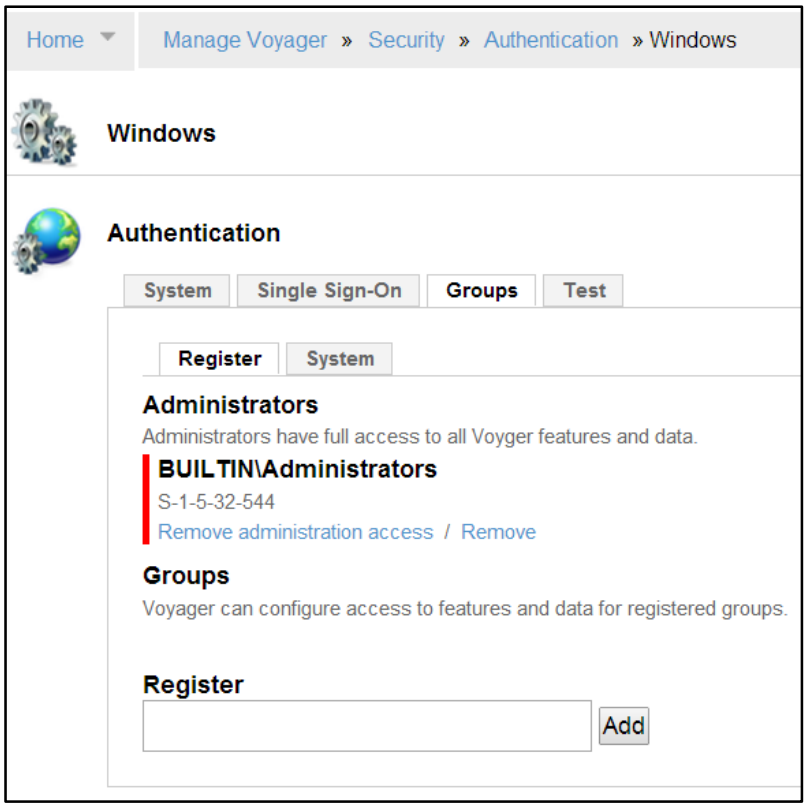
Voyager's configuration for Windows Authentication enables you to integrate Voyager into a Windows domain environment, providing a better search experience for your users. Single Sign-On uses domain credentials to log in to Voyager automatically.

Note: In order for Windows Authentication to be enabled, Voyager must be running as a service.

Windows Authentication Groups

To add a group using internal authentication:

1. Go to **Manage Voyager > Security > Authentication**
2. Select **Windows** and click **Configure**
3. Select the **Groups** tab



Single Sign-On

To set up Single Sign-On, first run the following command to register the Service Principal Name (SPN) for the Voyager server:

```
setspn -A HTTP/machine:port username
```

where `machine:port` is the machine and port the Voyager service is running on and `username` is the Voyager service account.

On the Single Sign-On tab, check **Enable Negotiate**. Click the SPN tab to test the configuration. You should see the Voyager server name under Registered ServicePrincipalNames.

If the Voyager server name does not appear under **RegisteredServicePrincipalNames** or the SPN tab shows an error, verify that Voyager is running as a service, and that the service account is on the same domain on the machine.

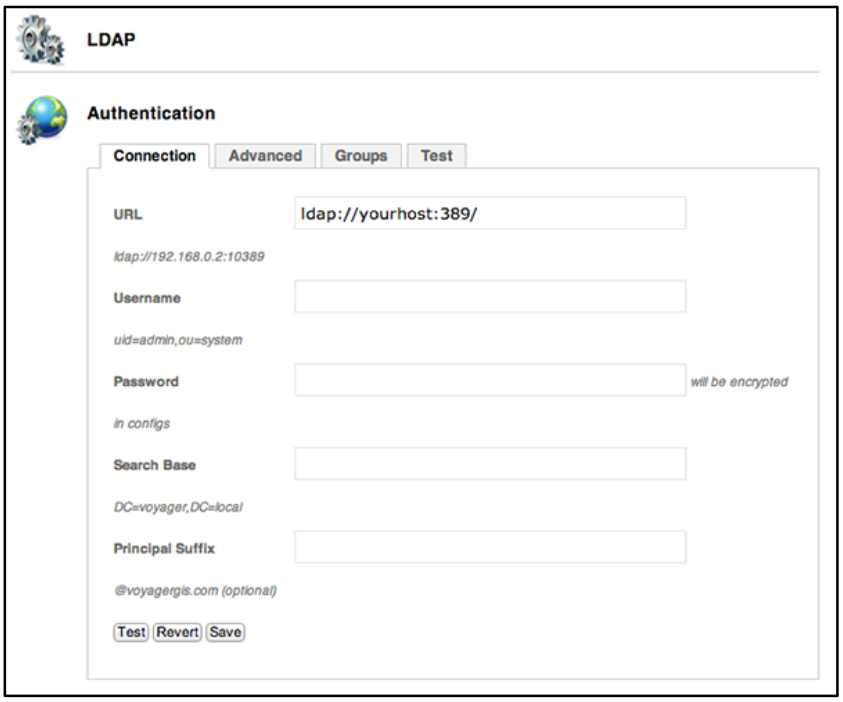
Re-run setspn \-A HTTP/machine:port username.

Check the SPN tab again. If the Voyager server name is still not present, check **Enable NTLM**.

LDAP Authentication

[**Manage Voyager > Security > Authentication > Security > LDAP**]

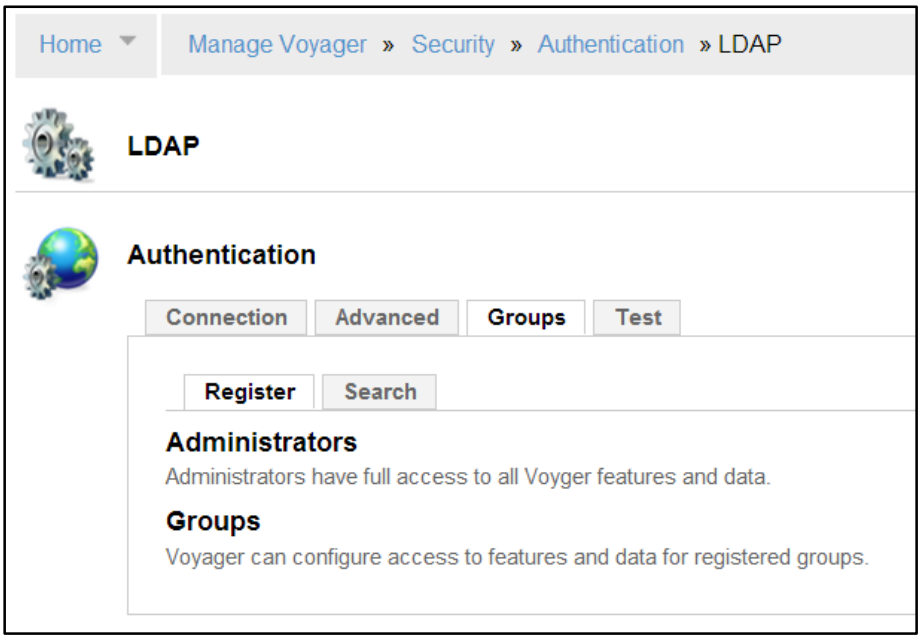
Voyager's LDAP connector enables you to integrate Voyager into existing security environments by mapping Voyager's role-based security to LDAP groups and user accounts.



LDAP Authentication Groups

To add a group using internal authentication:

1. Go to **Manage Voyager > Security > Authentication**
2. Select **LDAP** and click **Configure**
3. Select the **Groups** tab



Sessions

[**Manage Voyager > Security > Sessions**]

This page shows all of the current Voyager sessions.

Permissions

[**Manage Voyager > Security > Permissions**]

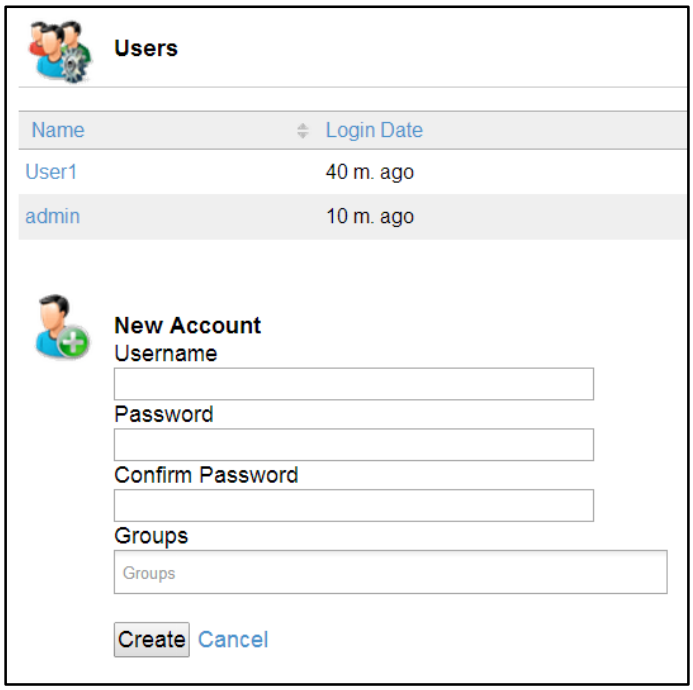
To configure permissions for anonymous and authenticated users, click the entry to allow or disallow that option.

	Anonymous*	Authenticated*
Use Voyager	✓	✓
Show Metadata	✓	✓
Show RSS Feeds	✓	✓
Use ArcMap Extension	✓	✓
Download	✓	✓
Upload	✓	✓
Configure View Settings	✓	✓
Federated Search	✓	✓
Show: Open With > ESRI ArcMap (.lyr)	✓	✓
Save Searches	✓	✓
Save Default Search	✓	✓
Allow Processing	✓	✓
Federated Processing	✓	✓
Refactor (Copy, Move, or Modify)	✓	✓
Register in ArcGIS Online	✓	✓
Enable Lists	✓	✓
Export Results	✓	✓
Configure Export Fields	✓	✓
Generate Report	✓	✓
Manage Voyager	✗	✗
Manage Security	✗	✗

Managing Users

[**Manage Voyager > Security > Users**]

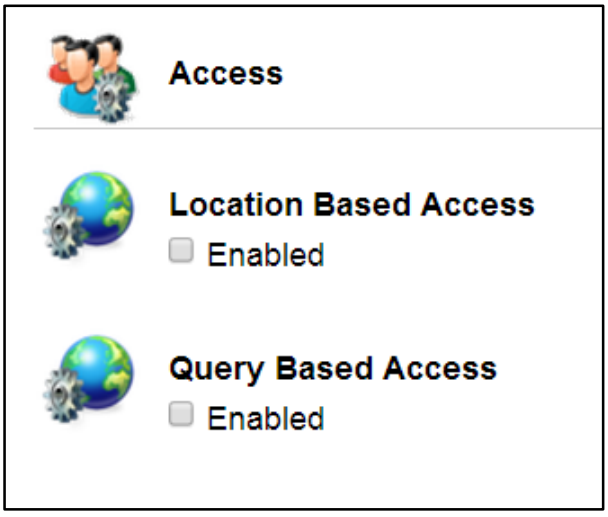
This page shows all Voyagers users and their login history. This is also where you can create new user accounts. To add a user, click **New Account** and fill in the **Username** and **Password** fields. You can also select a group for the new user.



Managing Access

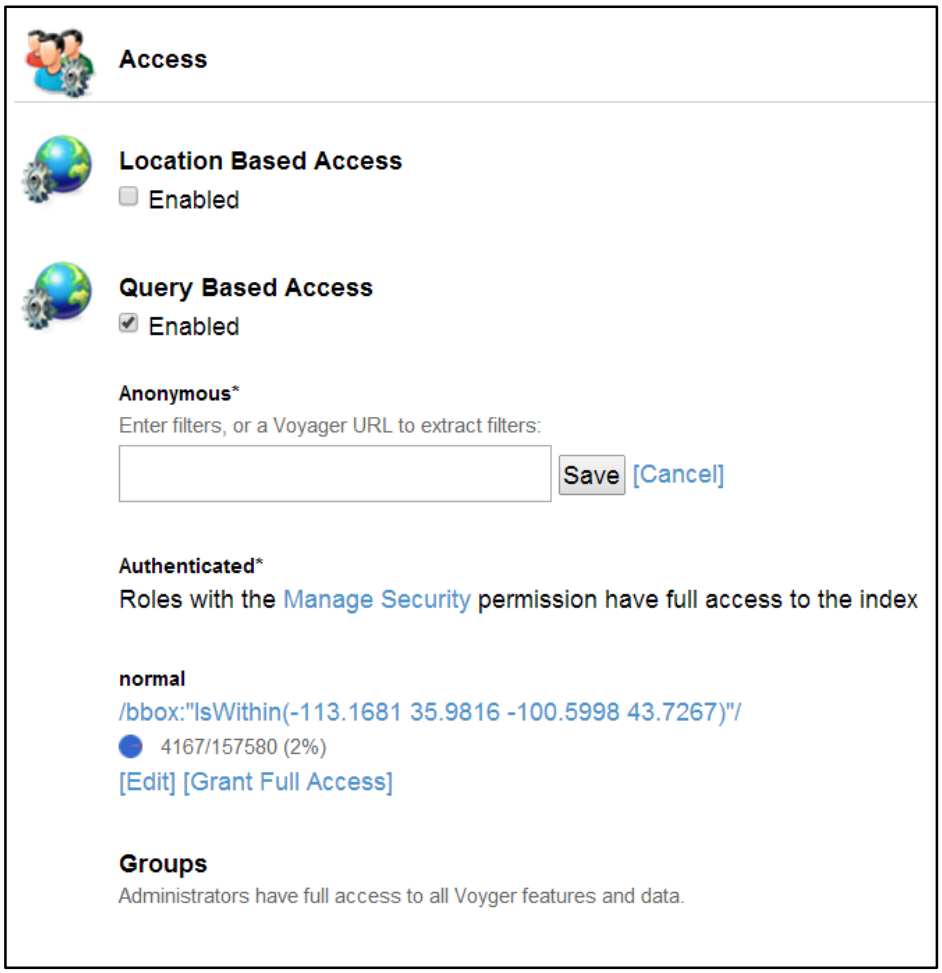
[**Manage Voyager > Security > Access**]

Use this option to configure location-based and query-based access restrictions.



Query-Based Access Restrictions

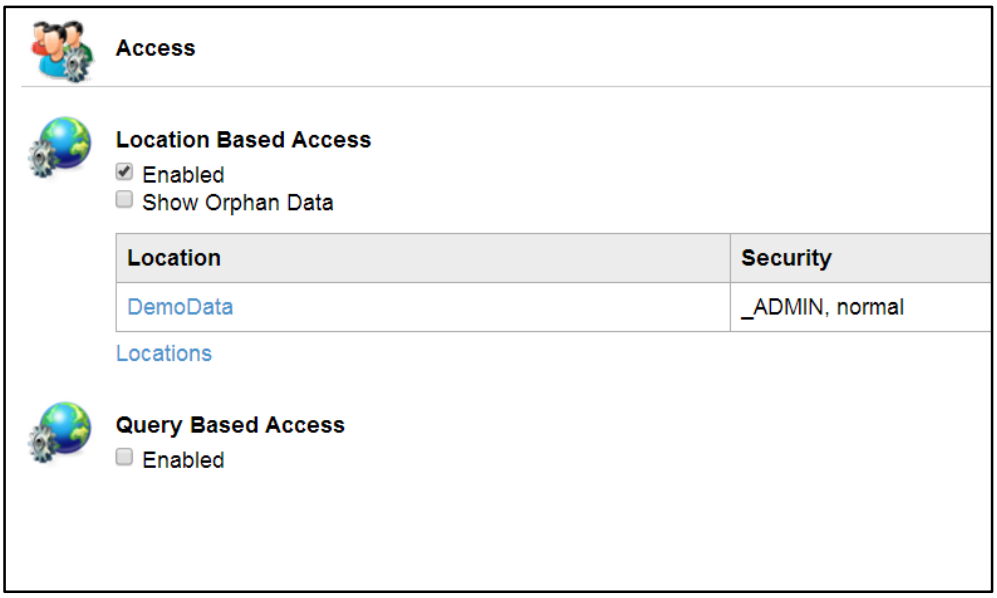
Query-based access allows an administrator to limit access based on any query. To specify a query, run a search with the appropriate keyword, spatial or location filter in another Voyager window, and then copy and paste the URL that results into this dialog.



Location-Based Access Restrictions

Location-based access allows an administrator to restrict access to the index based on the location of the content; it does not refer to the geospatial locations that the

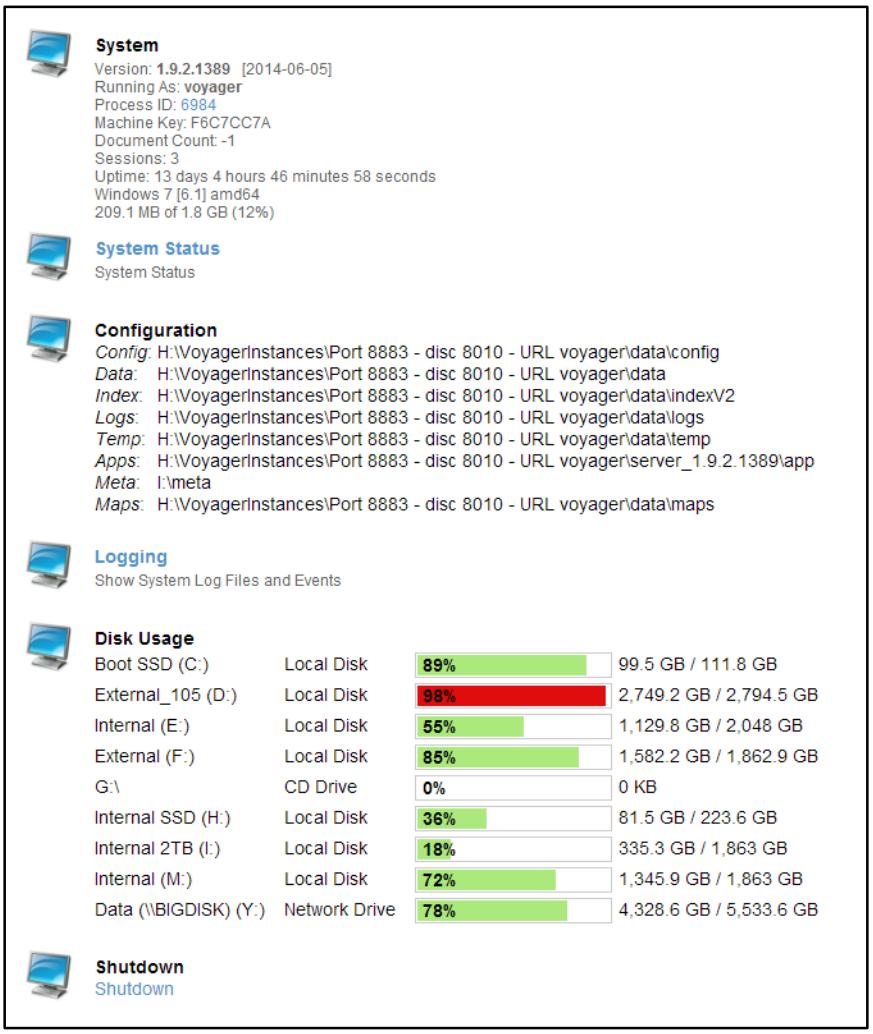
content may describe.



Managing System Settings

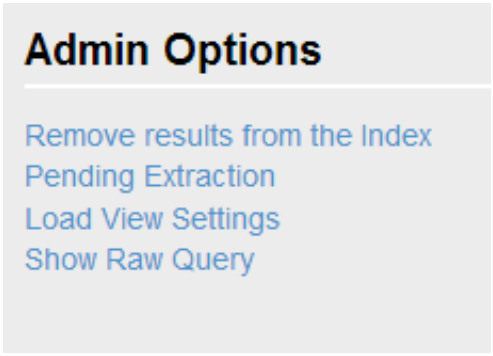
[**Manage Voyager > System**]

Displays System information, disk usage and System Logs and Events.



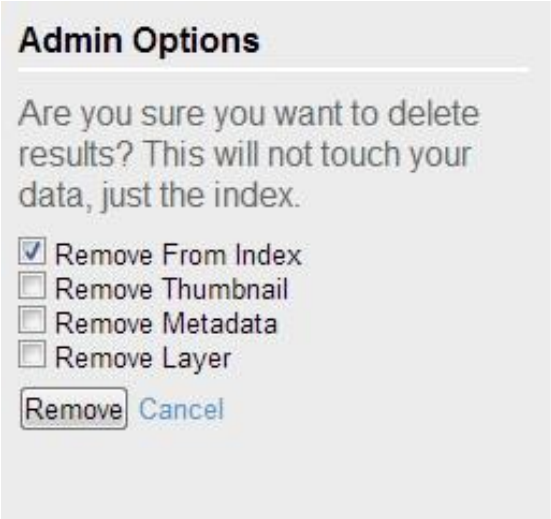
Special Admin Options

If you have administrator privileges, the following **Admin Options** are available at the bottom right of each summary or detail page.



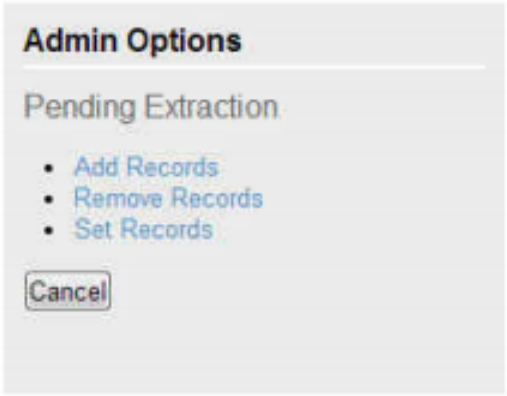
Remove Results from the Index

This removes search results from the Voyager catalog so that they will not appear in future searches. The source data itself is not affected. You have the option of also removing associated thumbnails, metadata, and layers.



Pending Extraction

Pending Extractions shows the status of pending extractions of information from the specified data locations.



Show Raw Query

This option shows the query that generated the current results page in either XML or JSON format. All available metadata fields are displayed for each record, in addition to the search parameters. The **debug** option also shows how the relevance score was calculated for each record.

Data Inventory and Assessment

Finding Duplicates

To use Voyager to find multiple copies of the same data, whether it's on your computer on somewhere else in the organization, go to **Saved Searches** and click **Duplicate Data**. Voyager displays the data in Grid view with a Content Hash and Schema Hash.

- The **Content Hash** indicates the same table schema AND the same data.
- The **Schema Hash** indicates the same table but DIFFERENT data.

Missing data links

You can use Voyager to find data links that are missing. Moving or deleting map documents (.mxd or .lyr) or data files often results in links being broken between the document and the data. Use the **Items with Broken Data** Saved Search to find which links are broken. This will generate a list of broken data links. The data that is missing will be highlighted in yellow.

By clicking on the Map link in the right hand column, you can now see the map documents and the broken links within them so that you can go and repoint to the correct layer.

Most commonly used data

You can use Voyager to commonly used data and see where it is used. Go to **Home** and **Saved searches**. Select **Most Commonly Used Data**.

This will generate a list of the most commonly used data sets in order of usage, but you can switch to a grid view. It is useful to see data that might be broken but that is commonly used so that it can be corrected.

Distributed Searching

Distributed Data Discovery

Use multiple computers to manage discovery process. This feature allows organizations with considerable processing needs to run tasks across many machines. Job scheduling can be used to control Distributed Data Discovery so that it only runs at times when computers would otherwise be idle.

Index Replication (Load Balancing)

To increase performance and availability, identical copies of the Voyager index can run in parallel.

Index Sharding

Sharding is useful for an extremely large number of documents (>50 million). Voyager can be configured to split the index across multiple computers while still maintaining the appearance and usability of a single index.

Web Application Clustering (Load-Balancing)

A load-balanced configuration is where Voyager runs in a clustered application server. In this configuration a group of servers running Voyager's application appears as if it were a single server. To balance load, Voyager distributes requests to different nodes within the server cluster, optimizing system performance and resulting in higher availability and scalability – a necessary option in large-scale enterprise, web-based configurations.

Hardware and Software Requirements

While data sets, hardware and network configurations can vary greatly, we recommend the following minimum requirements for most Voyager systems:

Hardware

- 8 CPU cores
- 10 GB memory
- 5 GB of disk space for the Voyager index (This should be on a very fast disk, such as where the OS is installed)
- 20+ GB disk space for thumbnails (This can be stored on Network Attached Storage or other ancillary hard drives)

Operating Systems

- A 32-bit edition of Windows XP or Windows Vista; or 32- or 64-bit Windows 7 or Windows Server.
- Windows 7 is preferred
- 64-bit operating systems are recommended

Web Browser

You can use Mozilla Firefox, Google Chrome or Internet Explorer.

ESRI Data Indexing Requirements

To index ESRI data formats you will need:

- A valid license of any ArcGIS™ 9.2, 9.3, 9.3.1 or 10.0 software product including ArcGIS Desktop™. Service packs are recommended.
- ArcGIS™ 9.2, 9.3, 9.3.1 .NET installation option must be enabled

The ArcGIS license will only be used during a Voyager process that requires it. For example, while indexing ESRI GIS data, a license will be checked out for each discovery thread that is indexing ESRI data. While running Voyager, ERSI licenses will also be used when running any ESRI-related Processing Task, for example **Clip and Ship** or **Mosaic**.

Microsoft Office 2007 is required to generate thumbnail images for Microsoft Word, Excel and Power Point documents. The software is not required for indexing.

ArcGIS Server is required when using Voyager to perform geoprocessing tasks in a server environment.

Virtual Machines

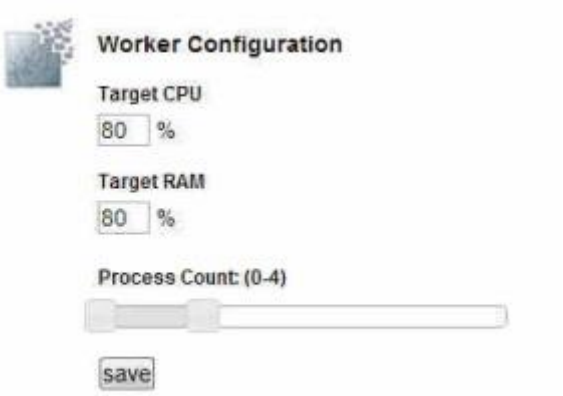
Virtual Machines are both supported and often recommended. Virtual Machines let you scale indexing resources (CPU cores) during initial indexing and then throttle them back later, when the system is in maintenance mode.

Additional Software

Voyager requires a license of ArcGIS Desktop versions 9.3.1, 10, 10.1,10.2 for users who want to index Esri proprietary formats such as .MXD, .LYR, and geodatabases. Users who do not want to index Esri proprietary formats do not require any additional software.

Additional Considerations

Voyager can leverage as many CPUs as are available. The more processor cores, the greater the number of indexing processes that can run simultaneously. The Worker Configuration settings can be found on the Discovery Configuration page. You can use these settings to tune Voyager so defined thresholds are not exceeded.



For average size indexes (100K-1M records), Voyager needs 2-3 GB of memory; for the web service, 1-2 GB per indexing process. The more processes, the more memory needed; however, this is also dependent on the types of data that you are indexing (some formats require more memory).

Every additional block of 1 million records will require about 512 MB - 1 GB of additional

system memory. This is not a linear equation. For systems indexing more than 30 million files, configuration recommendations vary. (Please inquire at support@voyagergis.com for more information)

Access to data

Voyager should have quick access to the data that it is indexing. When the data is on a local network attached storage device or the local machine, this is usually not a problem, but if Voyager is indexing data on remote disks or across a slow network (for example), indexing speed could be impaired.

Adding a Custom Task to Voyager

You can add custom tasks to Voyager using a Python script. There are several files you need to create before the custom task will appear in the **Processing Task Manager**.

Creating the info file

The first file is a JSON file that stores the parameters that your Python script will load later, and should be in the **Voyager\server_1.9\app\tasks\info** folder. The filename must end with **.info.json**, for example **my_function_py.info.json.**

Here is a sample .info.json file (copy_files.info.json), included with the Voyager installation:

```
{
  "name": "copy_files",
  "runner": "python",
  "categories": ["download"],
  "params": [
    {
      "type": "VoyagerResults",
      "name": "input_items",
      "required": true
    },
    {
      "type": "FolderLocation",
      "name": "target_folder",
      "required": true
    },
    {
      "type": "CheckBox",
      "name": "flatten_results"
    }
  ],
  "display":
  {
    "en":
    {
      "display": "Copy Files",
      "description": "Copies files to a target folder",
      "helpURL":
"https://github.com/voyagersearch/tasks/tree/master/docs#copy_files",
        "params":
        {
          "target_folder": {
            "display": "Target Folder",
            "description": "The folder where all files are copied. If it
does not exist, it will be created."
          },
          "flatten_results": {
            "display": "Flatten Results",
           "description": "Place all items in the root folder"
          }
        }
    }
  }
}
```

Creating the Python Script

The second file you need to create is the Python script itself, which should be saved in **Voyager\server_1.9\app\tasks\voyager_tasks**. It must have the same name as the info.json file you created earlier, in this example the info filename is **copy_files.info.json**, so the python filename is **copy_files.py**.

Here are the contents of **copy_files.py:**

```
# -*- coding: utf-8 -*-
# (C) Copyright 2014 Voyager Search
#
# Licensed under the Apache License, Version 2.0 (the "License");
# you may not use this file except in compliance with the License.
# You may obtain a copy of the License at
#
#  http://www.apache.org/licenses/LICENSE-2.0
#
# Unless required by applicable law or agreed to in writing, software
# distributed under the License is distributed on an "AS IS" BASIS,
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
# See the License for the specific language governing permissions and
# limitations under the License.
#from __future__ import unicode_literals
import os
import shutil
from voyager_tasks.utils import status
from voyager_tasks.utils import task_utils
def execute(request):
    """Copies files to a target folder.
    :param request: json as a dict.
    """
    parameters = request['params']
    input_items = task_utils.get_input_items(parameters)
    target_folder = task_utils.get_parameter_value(parameters, 'target_folder', 'value')
    flatten_results = task_utils.get_parameter_value(parameters, 'flatten_results',
'value')
    if not flatten_results:
        target_dirs = os.path.splitdrive(target_folder)[1]
        flatten_results = 'false'
    if not os.path.exists(request['folder']):
        os.makedirs(request['folder'])

    i = 1.
    copied = 0
    skipped = 0
    errors = 0
    file_count = len(input_items)
    shp_files = ('shp', 'shx', 'sbn', 'dbf', 'prj', 'cpg', 'shp.xml', 'dbf.xml')
    sdc_files = ('sdc', 'sdi', 'sdc.xml', 'sdc.prj')
    status_writer = status.Writer()

    status_writer.send_percent(0.0, _('Starting to process...'), 'copy_files')
    for src_file in input_items:
        try:
            if os.path.isfile(src_file) or src_file.endswith('.gdb'):
                if flatten_results == 'false':
                    # Maintain source file's folder structure.
                    copy_dirs = os.path.splitdrive(os.path.dirname(src_file))[1]
                    if not copy_dirs == target_dirs:
                        dst = target_folder + copy_dirs
                        if not os.path.exists(dst):
                            os.makedirs(dst)
                else:
                    if not os.path.exists(target_folder):
                        dst = target_folder
                        os.makedirs(dst)
```

```
                    else:
                        dst = target_folder
                if os.path.isfile(src_file):
                    if src_file.endswith('.shp'):
                        all_files = task_utils.list_files(src_file, shp_files)
                    elif src_file.endswith('.sdc'):
                        all_files = task_utils.list_files(src_file, sdc_files)
                    else:
                        all_files = [src_file]
                    for f in all_files:
                        shutil.copy2(f, dst)
                else:
                    shutil.copytree(src_file, os.path.join(dst,
os.path.basename(src_file)))
                status_writer.send_percent(i/file_count, _('Copied:
{0}').format(src_file), 'copy_files')
                copied += 1
            else:
                status_writer.send_percent(
                    i/file_count,
                    _('{0} is not a file or does no exist').format(src_file),
                    'copy_files'
                )
                skipped += 1
        except IOError as io_err:
            status_writer.send_percent(
                i/file_count, _('Skipped: {0}').format(src_file), 'copy_files')
            status_writer.send_status(_('FAIL: {0}').format(repr(io_err)))
            errors += 1
            pass

    try:
        shutil.copy2(os.path.join(os.path.dirname(__file__), 'supportfiles',
'_thumb.png'), request['folder'])
    except IOError:
        pass

    # Update state if necessary.
    if errors > 0 or skipped > 0:
        status_writer.send_state(status.STAT_WARNING, _('{0} results could not be
processed').format(skipped + errors))
    task_utils.report(os.path.join(request['folder'], '_report.json'), copied, skipped,
errors)
```

Note that the python file must have an **execute** statement or it will fail to run. To use your task, click the refresh button in the **Process > Task Manager window**.