Harvesting

Last updated: Fri, 15 Feb 2019 19:16:00 GMT

WorldCat Sync

CONTENTdm uses a tool called the Digital Collection Gateway to synchronize metadata from your collection to WorldCat.

To find out more about the details, see WorldCat Sync.

OAI Support

The Open Archives Initiative (OAI) supports a standard for disseminating metadata to other systems. By configuring this server setting, your CONTENTdm collections can be set up as an OAI Protocol for Metadata Harvesting (PMH) repository. The OAI-PMH repository responds to requests for your collection metadata from an OAI harvester service. OAI is sometimes used by library catalog systems to create references to your digital content.

If you enable OAI in CONTENTdm, the default setting allows all published, unrestricted collections to be available for harvesting. You can change the setting so that only specific collections are harvested. You can also opt to allow harvesting of page-level metadata in compound objects. Choosing to allow the harvest of pages is uncommon since most OAI harvesters are not designed to harvest transcripts or build full-text indexes.

For more information on OAI-PMH, visit www.openarchives.org.

OAI is disabled by default. To configure your OAI settings, you first must enable it.

To view and edit the OAI settings:

1. Open CONTENTdm Administration and select the server tab.
2. Click harvesting.
3. Review the current OAI settings.
4. Click the edit link next to OAI. The Edit OAI settings page opens.
5. Select Yes from dropdown menu to enable OAI.

Note: These settings apply to all collections included for harvesting.

**Base URL for OAI repository** will be in the format https://your.contentdm.website/oai/oai.php. This is the public-facing URL of your CONTENTdm site. If you have a custom domain or a domain at your institution, you should use that as your base URL and append /oai/oai.php to assemble the complete repository URL.

**Repository Name** is the name of the repository. Choose whatever name will be meaningful to the system that is harvesting your content. Your institution or library name is most common.

**Server Administration E-mail** is the email address of the server administrator. This is the person who would be contacted by the OAI harvester if they have questions about your content.

To enable compound object pages for harvesting, set **Enable compound object pages** to Yes.

- Enabling this setting allows OAI harvesters to harvest metadata for each page of the compound objects. Most CONTENTdm compound objects have little descriptive metadata at the page-level (Title only), so enabling this setting is typically done only when you want the page transcripts to be harvested for building a full-text index.

6. Click **save**. If no collections are specified in the **Collections** text box, all published, unrestricted collections on the server are available for harvesting.

7. If you want to specify which collections should be harvested, click the **edit** link next to **Collections**. The Edit OAI Collections window appears..
8. Follow instructions to **add** or **remove** collections for OAI harvesting.

**Stop list**

The stop list feature is no longer used by CONTENTdm because individual words are no longer hyperlinked in the end-user interface. Hyperlinking is now applied only to fields that have a controlled vocabulary. See [Items, Metadata](https://help.oclc.org/Metadata_Services/CONTENTdm/CONTENTdm_Administration/Server_Administration/020H) for more information.