Data Sources

Note

This information is not applicable to Primo VE environments. For more details on Primo VE configuration, see Primo VE.

Aleph

Primo harvests data from Aleph by utilizing Aleph’s publishing mechanism, which is available from version 16. The publishing mechanism extracts new and updated bibliographic records and related additional information (cross-references, holdings, items, and availability). For additional details about Aleph’s publishing mechanism, refer to the Aleph Publishing Mechanism, which includes information about the required setup for the Primo publishing mechanism.

DigiTool

Primo harvests data from DigiTool using the OAI harvesting mechanism. In DigiTool, you need to create a repository replication set, which is required for the OAI harvesting. For additional information, refer to the DigiTool Configuration Guide.

Within the repository replication, you need to create a set for Primo. The set defines the group of records that you want Primo to harvest. It is recommended to include only records with descriptive metadata in the set. If you create relations between records (the parent-child or manifestation type of relationship) within DigiTool, you may want to limit the set to include only the primary record. This is especially true if the secondary records include descriptive metadata that is not relevant to Primo. In terms of the types of metadata to include/retain, it is recommended to include descriptive and access rights metadata.

Use the PrimoTransformer when the records are replicated and then harvested. Note that if the digital object includes indexed full-text, this full-text will be embedded in the record and harvested by Primo. It is possible to add a flag limiting the size of the full-text that is added.

Note

Primo has separate rules for DigiTool-MARC21 and DigiTool-Dublin Core. If you have records with descriptive metadata in MARC21 and Dublin Core you should create two separate sets so that the records can be processed in two separate pipes.

MetaLib

Primo can harvest IRD records so that online databases can be included within Primo.
MetaLib Version 4 (from SP4) includes a special export option for Primo. For information on this option, refer to the How to Export KB Resources to Primo document under MetaLib version 4.x.

Note that MetaLib does not have an incremental export. It will always include the full DAT01 database. To ensure that any deleted resources are also deleted from Primo, the MetaLib pipe should be run with the Refresh Load option selected on the Pipe Definition page. A pipe run with this option will delete all of the records of the defined data source from the database before loading the newly harvested records.

**Rosetta**

Primo can harvest Rosetta content so that users can view digital content and share metadata. For configuration details, see the Rosetta - Primo Integration Guide.

**SFX**

Primo can harvest MARC21 records for e-journal and e-book resources from the SFX KnowledgeBase. To do this, export the records with active full-text journals and books. This may be required if these resources are not included in the ILS catalog.

SFX has a special export program for Primo (SFX2Primo), which can export either all records (for the initial harvesting) or new and updated records (for ongoing harvesting). The export utility creates the file in a dedicated directory, `exlibris/sfx_ver/sfx_version_3/<instance>/export/sfx2primo_export`, from which Primo FTP/SFTPs the files. The files are tarred and gzipped, as required by Primo.

**Note**

It is possible to include only e-journals, only e-books, or both (by selecting serials, monographs, or both). If you loaded e-journals from SFX into your catalog with MARCit! but do not have e-books in the catalog, create an export of e-books only.

The SFX export does not add an OAI header with the system ID. The system ID should be added by the publishing pipe in Primo. When you define the SFX data source, select Record does not include OAI header and choose the SFXOAI.xsl transformer. This ensures that the OAI header with the SFX system ID from the 090 tag is added to the records. If the record has been deleted from SFX (indicated by a `d` in pos. 5 of the leader), the program will add the deleted status.

For more information on the SFX harvesting functionality, see the SFX System Administration Guide.

**Unicorn**

Vanderbilt University has developed an extract procedure for Primo that includes new and updated bibliographic records and related additional information (cross-references, holdings, items, and availability). For details, contact Ex Libris support.
Voyager utilizes data extraction tools to customize and schedule results for the harvesting process. These tools extract new and updated bibliographic records and related additional information (such as cross-references, holdings, items, and availability), and create files containing the generated records that Primo accesses for harvesting and subsequently normalizes into PNX. Details for customizing these data extraction tools can be found in the Voyager-Primo Integration User's Guide.