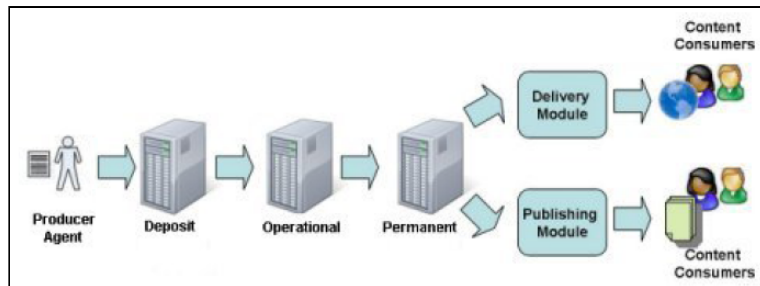


## The Rosetta System Architecture

### Understanding the Rosetta System Architecture

The following diagram illustrates the organization of the Rosetta system components and information flow.



#### The Rosetta System Architecture Overview

##### Note

Components can be managed on a single server or distributed among several servers. For more information, see [Storage Components](#).

The information flow consists of the following stages:

1. Producer Agents log on to the Rosetta system.
2. Producer Agents upload files through the deposit server.
3. The Rosetta system moves the Producer Agents' content to the operational server, which processes content through the SIP Processing module.
4. During SIP processing, staff users review the content that Producer Agents deposited. They decide whether to approve the content, return it to Producer Agents for repairing issues ("reject" it), or decline it due to content issues.
5. Approved content is moved to the permanent repository. Returned and declined content is sent back to Producer Agents.
6. Both the Delivery Module and Publishing Module deliver the content from Rosetta through the interface to the content consumers.

## Authentication

Rosetta can authenticate users locally with Rosetta or externally using the SAML protocol. SAML is used for external authentication for users managed by an IdP. SAML is supported as of Rosetta version 5.2. For more information, see [SAML User Authentication](#).

---

## Storage Components

The following components store the content that Producer Agents deposit:

- [Deposit Server](#)
- [Operational Server](#)
- [Permanent Repository](#)

Each storage component is characterized by the data it stores and the users who can work with this data.

This guide refers to each component as a singular server. In fact, your library may store all components on one server or use a single server for each type of storage. References to a “server” such as an “operational server,” therefore, may refer to a portion of a server or an entire server in a group of three or more servers.

For a description of server topology, see chapter 1 of the *Rosetta System Administration Guide*.

---

### Deposit Server

The deposit server receives content from Producer Agents. Producer Agents can upload files of various formats and provide descriptive information about their content (such as title, author, and creation date).

The types and number of files, as well as the types of descriptive information that Producer Agents must provide, are defined by staff users. (For more information, see [Material Flows](#).)

Producer Agents use a Web-based wizard to deposit content.

#### *Deposit Server Data*

To logically organize content that Producer Agents deposit, the system groups Producer Agent files into deposit activities on the deposit server. (A deposit activity contains files that a Producer Agent has provided as one deposit.) In addition, deposit activities contain descriptive information (such as title, author, and creation date) that Producer Agents have provided about the content.

The deposit server stores the following types of deposit activities:

- Deposit activities that Producer Agents have saved as drafts for future deposits. Producer Agents can edit the draft deposit activities by adding new files, replacing the files, or editing descriptive information. When the draft deposit activity is ready, Producer Agents submit it for deposit.
- Deposit activities that staff users have returned to Producer Agents after review. Producer Agents can repair the issues and resubmit the deposit activity.
- Deposit activities that staff users have declined after review. Producer Agents cannot resubmit declined deposit activities.

Producer Agents manage their deposit activities through the Web-based interface.

From the deposit activities, the system creates SIPs and passes them on to the operational server.

#### *Deposit Server Users*

The following types of users work with the deposit server:

- Content providers:
  - **Producers** are responsible for providing content. (For more information, see [Producer Types](#).)

- **Producer Agents** are responsible for depositing content. They self-register through the Web-based interface. (For more information, see [Content Providers](#).)
- Staff users:
  - **Deposit Managers** are responsible for configuring generic settings for Producers. The Producer Agents associated with these Producers use these settings when they deposit content. For example, Deposit Managers define the types and maximum size of files that associated Producer Agents can deposit. Deposit Managers are assigned and registered by a Back Office Administrator. (For more information, see the Deposit Managers section in the *Rosetta Staff User's guide*.)
  - **Negotiators** are responsible for tailoring the generic settings to the needs of specific Producers. For example, Negotiators can allow associated Producer Agents to deposit additional types of files.

Negotiators are assigned and registered by a Back Office Administrator. (For more information, see Negotiators in the *Rosetta Staff User's guide*.)

---

## Operational Server

After a deposit activity is submitted by a Producer Agent, the deposit server packages it as a SIP and notifies the operational server that this SIP is ready for processing. The SIP then goes through the automated phase of validation (validation stack), and then it becomes available to staff users. Staff users review the content and decide whether it needs to be approved, returned to the Producer Agent, or declined.

To access and review the content, staff users work with the Web-based interface.

### Creating a SIP

To enable staff users to work with the Producer Agents' content, the Rosetta system moves the deposit activity to the operational server and converts each deposit activity into a Submission Information Package (SIP). (For more information, see [Submission Information Packages \(SIPs\)](#).)

## Operational Server Users

The following types of staff users work with data stored on the operational server:

- **Assessors, Arrangers, and Approvers** are responsible for reviewing SIPs and deciding whether a SIP must be approved, rejected, or declined. Assessors, Arrangers, and Approvers are assigned and registered by a Back Office Administrator. (For more information, see Assessors, Arrangers, and Approvers in the *Rosetta Staff User's Guide*.)
- **Technical Analysts** are responsible for repairing technical issues that may occur with the SIPs. For example, Technical Analysts can manually assign a format to a file that couldn't be associated automatically. Technical Analysts are assigned and registered by a Back Office Administrator. (For more information, see Technical Analysts in the *Rosetta Staff User's Guide*.)

---

## Permanent Repository

After a SIP is approved by staff users, the Rosetta system moves the intellectual entities (IEs) to the permanent repository. In the permanent repository, IEs are no longer grouped together as SIPs, though they retain, individually (as IEs), the metadata from the SIP.

From the permanent repository, the content can be delivered to content consumers through the Web and other channels.

## Permanent Repository Data

The permanent repository is intended to store Producer Agent content that was approved by staff users for permanent preservation. As a result, IEs that are stored in the permanent repository cannot be updated, deleted, or rearranged.

When an IE must be changed (for example, its metadata or its format requires updating), the Rosetta system moves it back to the operational server. When the updating process is finished, the system returns the IE to the permanent repository and stores it as a new version of the IE.

## Permanent Repository Users

The following types of users work with data stored in the permanent repository:

- **Editors** are responsible for managing content that was deposited by Producer Agents and approved by staff users for storage in the Rosetta system. (For more information, see Editors in the *Rosetta Staff User's Guide*.)
- **Content consumers** can search and view content in read-only mode. They access repository content through a public interface such as their library's online public access catalog (OPAC). The content that is available to them is defined by staff users (who define access rights options) and Producer Agents (who select specific options from those configured by staff users).

---

## OAI Provider

Rosetta uses the Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH) for publishing IE descriptions to external systems. These systems (such as Primo) harvest Rosetta for records by calling the OAI Data Provider component.

This component is an integral part of Rosetta and cannot be configured or deactivated from the UI.

The OAI-PMH requests are expressed as HTTP requests. The base URL specifies the Internet host and port. The URL continues with a list of keyword arguments that take the form of `key=value` pairs. Arguments are separated by ampersands (&).

The OAI-PMH `key=value` pairs must use arguments that are supported by Rosetta.

The arguments supported are:

- `Verb = ListRecords`. This argument is mandatory.
- `metadataPrefix` = This argument is mandatory. The only supported values are `oai_dc` and `xepicur`.
- `set = set_spec` as it is defined in the Publishing Profile.
- `from` and `until` are optional.

Request example:

```
http:// <rosetta-server>/oaiprovider/  
request?verb=ListRecords&set=books&metadataPrefix=oai_dc&from=2010-01-01T08:05:04Z&until=2010-10-
```

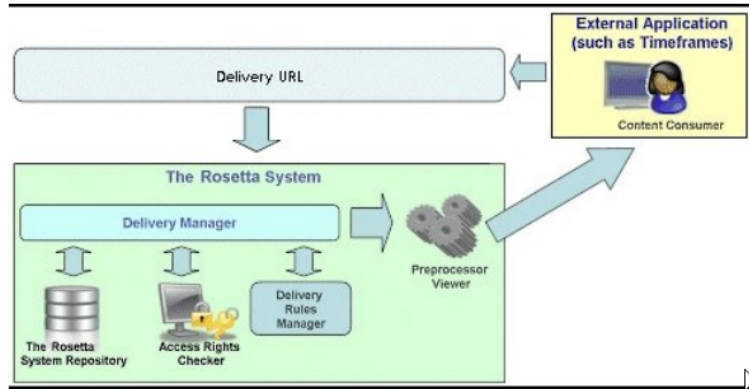
---

## Delivery

The Rosetta system enables certain users to view content objects (such as intellectual entities, representations, and files) that are stored in the Rosetta system. These objects can be viewed by staff users (for example, Assessors can view the content that Producer Agents deposited) and external users (for example, a reader who has a subscription to the library).

For the purpose of this guide, these users are called content consumers.

The following diagram illustrates the organization of the components that enable content delivery.



### Delivery Information Flow

The delivery flow assumes the integration of an external viewer application such as a library's OPAC with Rosetta's delivery system. It consists of the following stages:

1. A content consumer uses an external application to request a content object from the Rosetta system.
2. The external application sends the request to the Delivery Manager module of the Rosetta system by using the Delivery URL. The requests are expressed as HTTP requests. The base URL specifies the Internet host and port. The URL continues with the PID of the requested IE or Representation (for example, `http:// <rosetta-server>/delivery/DeliveryManagerServlet?dps_pid=IE1077`). For more details about the Delivery URL and optional parameters, see the *Rosetta Configuration Guide*.
3. The Delivery Manager retrieves the content from the repository.
4. The Access Rights Checker determines whether the content consumer has the appropriate privileges to view the requested content object.
5. If the content consumer has the relevant access rights, the Delivery Rules Manager verifies the input parameters of the content object (for example, whether the object is an IE, representation, or file), and determines which viewing profiles, viewer preprocessor, and viewer must be used to display the content object.
6. The content object is forwarded to the Rosetta system viewer preprocessor, which then prepares the content for display in the viewer used by the external application.
7. When the viewer preprocessor finishes processing the content object, the Rosetta system sends the object to the external application viewer.
8. The viewer displays the content object to the content consumer.

For more information about delivery, see the chapter on Configuring Delivery in the *Rosetta Configuration Guide*.