Configure ProQuest Access

Purpose

This page describes how to configure access to ProQuest in Circulation Manager.

Prerequisites

1. Ensure that the organisation that wants to use ProQuest as a vendor can authenticate its patrons using SAML.
2. Ensure that the organisation's IdP can send back to Circulation Manager a unique patron ID as a SAML attribute.
3. Ensure that you have machines with static IP addresses that ProQuest can whitelist.

Configuring Circulation Manager

Configuration settings

The table below contains a description of different ProQuest integration's configuration settings:

Parameter	Mandatory	Default	Description
Data source name	Yes	**ProQuest**	The name of the Circulation Manager's data source used to store information about ProQuest books.
BooksCatalog Service's URL	Yes		The URL of **BooksCatalog** service used to download the ProQuest OPDS 2.0 feed and described on ProQuest API page. Please note that the **BooksCatalog** service is accessible only from whitelisted IP addresses. Please refer to ProQuest API for more information.
PartnerAuthToken Service's URL	Yes		The URL of **PartnerAuthToken** service used to create a ProQuest JWT bearer token required to authenticate in other ProQuest services. and described on ProQuest API page
DownloadLink Service's URL	Yes		The URL of **DownloadLink** service used to download ProQuest books and described on ProQuest API page.
ProQuest JWT token's expiration timeout	No	3600	Determines how long in seconds can a ProQuest JWT token be valid.
List of SAML attributes containing an affiliation ID	No	eduPersonPrincipalName, eduPersonScopedAffiliation	ProQuest integration assumes that the SAML provider is used for authentication. ProQuest JWT bearer tokens required by the most ProQuest API services are created based on the affiliation ID - SAML attribute uniquely identifying the patron.This setting determines what attributes the ProQuest integration will use to look for affiliation IDs. The ProQuest integration will investigate the specified attributes sequentially and will take the first non-empty value.
Test SAML affiliation ID	No	1	This configuration setting is really useful for testing ProQuest API without setting up a SAML authentication provider. If this configuration setting is set, then the ProQuest integration will be using its value in the case when it cannot find a SAML affiliation ID. Please contact ProQuest before using it.
Default audience	No		If ProQuest does not specify the target audience for their books, assume the books have this target audience.
Feed page's size	No	5,000	This value determines how many publications will be on a single page fetched from the **BooksCatalog** service. Please note that this value also affects how many publications will be committed in a single database transaction by Circulation Manager. Making this value too big may lead to different database issues.
HTTP proxy's URL	No		URL of the proxy handling HTTP traffic. Proxies can be used for testing when trying to access ProQuest API from machines having non-whitelisted IP addresses. Please find more information below.
HTTPS proxy's URL	No		URL of the proxy handling HTTPS traffic. Proxies can be used for testing when trying to access ProQuest API from machines having non-whitelisted IP addresses. Please find more information below.

Proxies

As it was mentioned above it might be useful to use a proxy to test the SAML integration from the local machine which has an IP address non whitelisted by ProQuest. To do so you need to use a SOCKS proxy:

```
ssh -D 8888 -f -C -q -N <FQDN or IP address of the machine with a whitelisted IP address>
```

After you activated a SOCKS proxy, you'll need to create an HTTP proxy based on it:

```
npm install -g http-proxy-to-socks
hpts -s 127.0.0.1:8888 -p 8080
```

After this, you can use the following values for proxy settings:

- **HTTP proxy's URL = http://localhost:8080**
- **HTTPS proxy's URL =** https://localhost:8080

Depending on your system configuration you might face a weird **SSL: WRONG_VERSION_NUMBER** error. In this case you can try to use a SOCKS proxy:

- **HTTPS proxy's URL =** socks5://localhost:8888

Creating a ProQuest collection in Circulation Manager

1. Create a new collection with **ProQuest** protocol.
2. Set up all the mandatory configuration settings mentioned in the table above:

3. Set up optional configuration settings if required:

Optional Fields ⌃

ProQuest JWT token's expiration timeout

3600

(Optional) Determines how long in seconds can a ProQuest JWT token be valid.

List of SAML attributes containing an affiliation ID

ProQuest integration assumes that the SAML provider is used for authentication. ProQuest JWT bearer tokens required by the most ProQuest API services are created based on the affiliation ID - SAML attribute uniquely identifying the patron.This setting determines what attributes the ProQuest integration will use to look for affiliation IDs. The ProQuest integration will investigate the specified attributes sequentially and will take the first non-empty value.

eduPersonPrincipalName   Delete 🗑

eduPersonScopedAffiliatior   Delete 🗑

uid ⬍   Add
(Optional)

Test SAML affiliation ID

1

(Optional) Test SAML affiliation ID used for testing ProQuest API. Please contact ProQuest before using it.

Default audience

Adult ⬍

If ProQuest does not specify the target audience for their books, assume the books have this target audience.

Feed page's size

5000

ProQuest Import Monitor

Please note that the import of ProQuest collections is done on a scheduled basis, the same way it's done for OPDS 1.x/OPDS 2.x collections.

However, the ProQuest import process is a very computationally expensive operation due to the inefficient ProQuest feed's internal structure and **proquest_import_monitor** is set up to run only twice a week:

- Tuesday 2 AM EST (Monday 11 PM PST).
- Friday 2 AM EST (Thursday 11 PM PST). This run uses an additional **--process-removals** flag which forces **proquest_import_monitor** to remove from Circulation Manager's catalog items that are no longer present in the ProQuest feed (it's not enabled by default).

The schedule above is subject to change.