

EQUELLA®

Harvester Profile Configuration Guide

Version 6.2

Document History

Document No.	Reviewed		Finalised		Published	
1	2/12/2013		2/12/2013		2/12/2013	

December 2013 edition.

Information in this document may change without notice. EQUELLA® and its accompanying documentation are furnished under a non-disclosure, evaluation agreement or licence agreement. Copying, storing, transmitting, or otherwise reproducing the software or this document in any form without written permission from Pearson is strictly forbidden.

All products, other than EQUELLA®, named in this document are the property of their respective owners.

Property of:

Pearson
Level 1, 11 Elizabeth Street
Hobart, Australia, 7000

Copyright © 2013 Pearson

Contact documentation@equella.com for matters relating to documentation.

Table of Contents

Harvesting guide	4
Harvester Profile tool	4
Configure a harvester profile	4
Harvester types.....	7
Details tab	8
EQUELLA Harvester.....	9
LORAX Harvester	10
MEX Harvester	10
OAI Harvester	11
SHEX Harvester.....	12
Actions tab	13
Access Control tab	15
Harvester profiles scheduled task	17
Harvester system settings	19
Contact Client Support	20

Harvesting guide

Content from third-party sources can be obtained by using EQUELLA's harvesting tool that is used to collect learning objects and resources from third-party institutions and upload them automatically to the EQUELLA repository. Content can be obtained from The Learning Federation using the LORAX protocol, while content from other institutions can be gathered with the OAI (Open Archives Initiative) metadata harvesting protocol.

Content contained in the EQUELLA Digital Repository can also be accessed by using the OAI protocol.

The purpose of this guide is to provide system administrators with an understanding of the EQUELLA Harvester Profile tool to enable successful management of harvesting.

Please note that this guide has been developed to best reflect the full capabilities of EQUELLA and as such may differ in appearance to your own installation.

Harvester Profile tool

The EQUELLA harvester allows content from third-party sources to be collected and uploaded to the EQUELLA Digital Repository. Content can be obtained from The Learning Federation (TLF) using the Learning Object Repository Access and eXchange (LORAX) protocol, while metadata can be gathered with the Open Archives Initiative (OAI) metadata harvesting protocol. Content from other EQUELLA Digital Repositories can also be collected.

The Harvester Profile:

- provides harvesting of metadata and resources,
- allows harvesting of other EQUELLA repositories, enabling community repositories to be created with sizable pools of content,
- allows multiple Harvester Profiles to be created and managed via the Administration Console,
- can be run as a scheduled task.

Configure a harvester profile

Harvester profiles are created and modified using the **Harvester Profiles** tool available in the Administration Console. An example is shown in Figure 2.

To access EQUELLA and open the Administration Console

1. Open a browser and enter your EQUELLA URL (e.g. '<http://equella.myequellainstitution.edu>').
2. Log in to EQUELLA as an administrator user to display the EQUELLA dashboard.
3. Click the **Settings** button to display the Settings page, then select **Administration console** from the categories list. An example is shown in Figure 1.

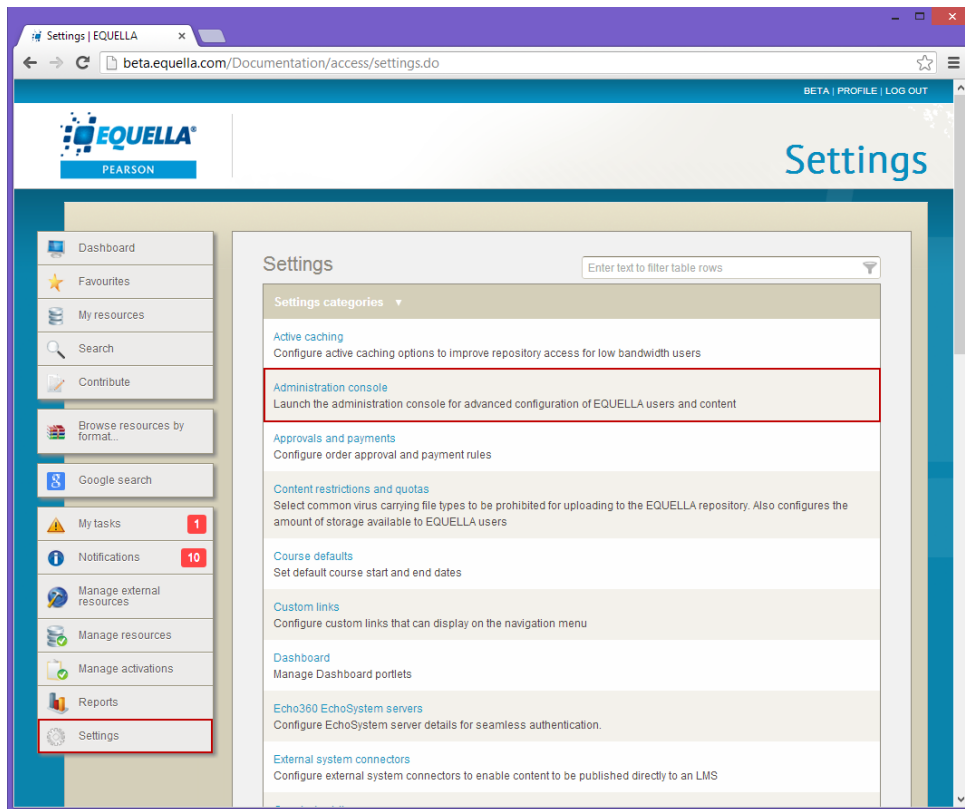


Figure 1 Open the Administration console from the Settings page

4. Select the **Harvester Profiles** option to display the **Harvester Profiles** pane that contains a list of all previously configured profiles. An example is shown in Figure 2.

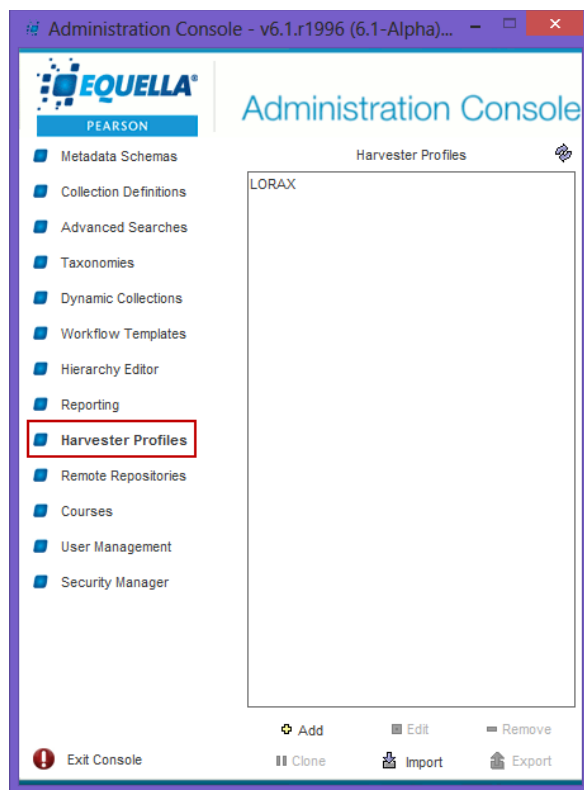


Figure 2 Administration Console—Harvester Profiles pane


Administration Console elements

When a tool has been selected the **Add**, **Edit**, **Remove**, **Clone**, **Import** and **Export** buttons are enabled. (Providing the correct privileges have been granted, refer to the *EQUELLA Security Administration Guide* for further information.)

Add

Click  **Add** to display the **Harvester Profiles Editor** and create a new profile.

Edit

Select a profile then click  **Edit** (or double-click on a profile) to display the **Harvester Profiles Editor** and modify the selected profile.

If the selected profile is being edited by another user a **Locked** dialog is displayed as shown in Figure 3, providing the following options:

- **Open as read-only**—disables the **Save** button so that no changes can be made.
- **Forcefully unlock**—unlocks the entity from the previous owner. Editing is possible but changes by other users may be lost.
- **Do not open**—the profile configuration screen will not be opened.

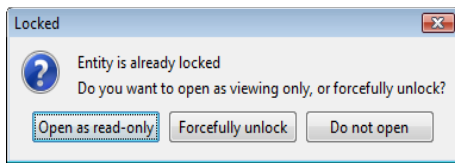




Figure 3 Locked entity dialog


Remove

Click the  **Remove** button to delete the selected profile.

Clone

Click the  **Clone** button to copy an existing profile. The cloned profile can then be modified by selecting the **Edit** button.

Import


Click the  **Import** button to import a harvester profile that has previously been exported from the Administration Console. A warning dialog is displayed as shown in Figure 4.

Note: Only ZIP files that have been exported from the Administration Console should be imported—other ZIP files may adversely affect the stability of the system.



Figure 4 Import HarvesterProfile warning dialog

Export

Click the  **Export** button to export a profile as a ZIP file.


Harvester types

Each Harvester profile type provides options for harvesting resources or metadata from different repository types:

- EQUILLA Harvester
- LORAX Harvester
- MEX Harvester
- OAI Harvester
- SHEX Harvester

Add a harvester profile

To add a profile

1. Click  **Add** on the Administration Console to display a **Harvester Types** dialog listing the available harvester types, an example is shown in Figure 5.

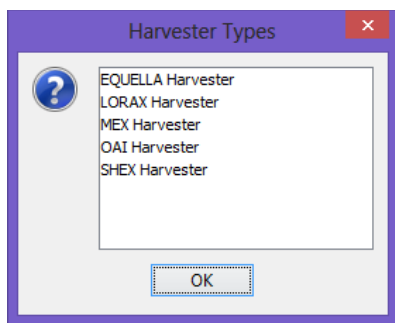



Figure 5 Harvester types dialog

2. Select a harvester type from the list.
3. Click  to display the **Harvester Profile Editor**.

The **Harvester Profile Editor** provides an interface for harvester creation with the following tabs:

- **Details**—set general information for the creation of the harvester profile.
- **Actions**—perform actions on the selected profile.
- **Access Control**—define privileges for access to the selected profile.

Each tab is described in the following sections.

Details tab

The **Details** page displays the harvester's general details such as name, enabled status, collection to harvest to and import transformation. These elements are common to all harvester profile types and are displayed in Figure 6.

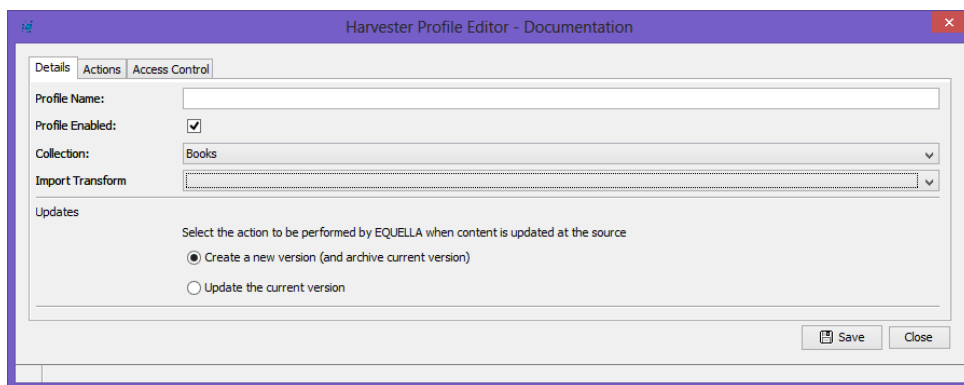


Figure 6 Harvester Profile Editor—Details page

Each type of harvester has other various configurable requirements. The **Harvester Profile Editor—Details** page contains configurable elements specific to the selected harvester type as well as the common elements shown in the figure above.

The common **Details** page elements are:

- **Profile Name**—enter the unique display name of the profile.
- **Profile Enabled**—is checked by default. Uncheck the box to disable the profile and stop harvesting.

- **Collection**—displays a list of collections where the harvested resources can be contributed.
- **Import Transform**—displays a list of XSLT for transforming the harvested resources metadata to that of the contributed collection. The list comprises XSL transformations that have been added to the metadata schema for the associated collection.
- **Updates** – select how updates occurring to a resource at the source affects the harvested resource in EQUELLA. The options are:
 - **Create a new version (and archive the current version)** – selecting this option creates a new version of the resource, while archiving the old.
 - **Update the current version** – select this option to update the resource with the changes from the source, but keep the current version number (similar to an edit).

The harvester type specific elements are described in the following sections:

EQUELLA Harvester

The EQUELLA type allows harvesting of resources and attachments from other EQUELLA repositories.

The screenshot shows the 'Harvester Profile Editor - Documentation - EQUELLA Documentation' window. It has three tabs: 'Details', 'Actions', and 'Access Control'. The 'Details' tab is active and contains the following sections:

- Profile Name:** EQUELLA Documentation
- Profile Enabled:** ☒
- Collection:** Books (dropdown menu)
- Import Transform:** MARCXML to Book (dropdown menu)
- Updates:**

Select the action to be performed by EQUELLA when content is updated at the source

☐ Create a new version (and archive current version)

☒ Update the current version
- EQUELLA Settings:**

Server: http://beta.equilla.com/Documentation/

Username: beta

Password: [masked with dots]

Remote Collection: Books (dropdown menu) Get Collections

Only harvest live resources: ☒

At the bottom right, there are 'Save' and 'Close' buttons.

Figure 7 EQUELLA—Details page

The configurable **EQUELLA Settings** include:

- **Server**—enter the URL for the EQUELLA institution to be harvested.
- **Username/ Password**—enter the username and password used to access the repository (if authentication is required).
- **Remote Collection**—select a collection to be harvested from the drop-down list. Select the **Get Collections** button to display a list of collections names.
- **Only harvest live resources**—is checked by default. Uncheck the box to include non-live resources in the harvest.

LORAX Harvester

The LORAX protocol is used to harvest TLF content.

Harvester Profile Editor - Documentation - LORAX

Details Actions Access Control

Profile Name: LORAX

Profile Enabled: ☒

Collection: Books

Import Transform: MODS to Book

Updates

Select the action to be performed by EQUELLA when content is updated at the source

☒ Create a new version (and archive current version)

☐ Update the current version

LORAX Settings

Username: equellaUser

Password: ••••••

Only check Live EQUELLA resources ☒

Harvest Learning Objects ☒

Harvest Resources ☐

Save Close

Figure 8 LORAX—Details page

The configurable **LORAX Settings** are:

- **Username/ Password**—enter the username and password used to access the repository.
- **Only check Live EQUELLA resources**—is selected by default and scans your EQUELLA Institution for *Live* resources to ensure that duplicates are not harvested. Uncheck the box to include non-live resources (e.g. *Draft*, *Archived*) in the checking process.
- **Harvest Learning Objects**—is selected by default. Uncheck the box to exclude learning objects in the harvest.
- **Harvest Resources**—is un-selected by default. Select the box to include resources in the harvest.

MEX Harvester

MEX is an Australian protocol used to harvest free content provided under the National Digital Learning Resource Network (NDLRN).

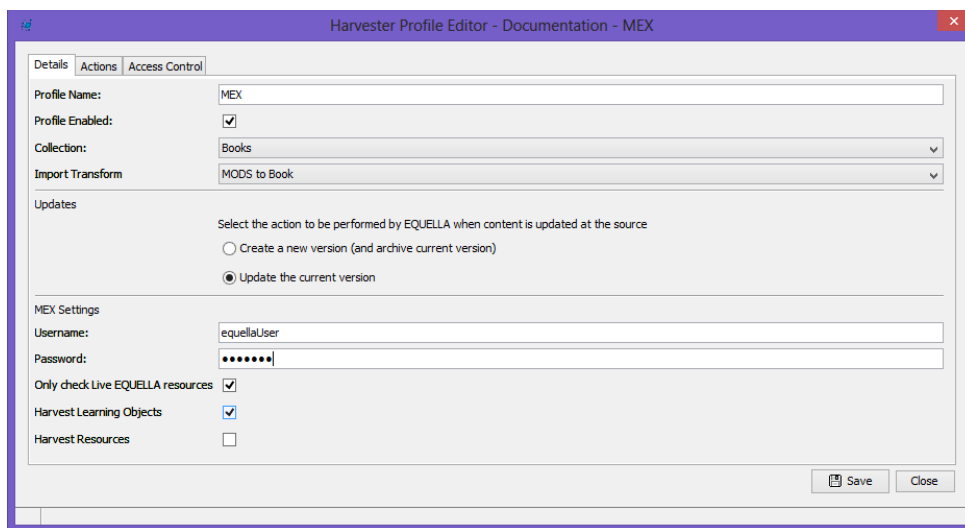


Figure 9 MEX - details page

The configurable **MEX Settings** are:

- **Username/ Password**—enter the username and password used to access the repository.
- **Only check Live EQUELLA resources**—is selected by default and scans your EQUELLA Institution for *Live* resources to ensure that duplicates are not harvested. Uncheck the box to include non-live resources (e.g. *Draft*, *Archived*) in the checking process.
- **Harvest Learning Objects**—is selected by default. Uncheck the box to exclude learning objects in the harvest.
- **Harvest Resources**—is un-selected by default. Select the box to include resources in the harvest.

OAI Harvester

OAI is a protocol that is used to harvest metadata from third-party sources.

Figure 10 OAI—Details page

The configurable **OAI Settings** are:

- **Server**—enter the URL for the OAI repository to be harvested.
- **Metadata Format**—select the format from the drop-down list that the metadata is retrieved in. Select the **Get Metadata Formats** button to display a list of formats.
- **Set**—select the OAI record set to be harvested from the drop-down list. Select the **Retrieve Sets** button to display a list of sets.

SHEX Harvester

SHEX is an Australian protocol used to harvest free content provided under the National Digital Learning Resource Network (NDLRN).

Figure 11 SHEX - details page

The configurable **SHEX Settings** are:

- **Username/ Password**—enter the username and password used to access the repository.
- **Only check Live EQUELLA resources**—is selected by default and scans your EQUELLA Institution for *Live* resources to ensure that duplicates are not harvested. Uncheck the box to include non-live resources (e.g. *Draft*, *Archived*) in the checking process.
- **Harvest Learning Objects**—is selected by default. Uncheck the box to exclude learning objects in the harvest.
- **Harvest Resources**—is un-selected by default. Select the box to include resources in the harvest.

Actions tab

The **Actions** page provides a flexible method for scheduling and running the harvest. An example is shown in Figure 12.

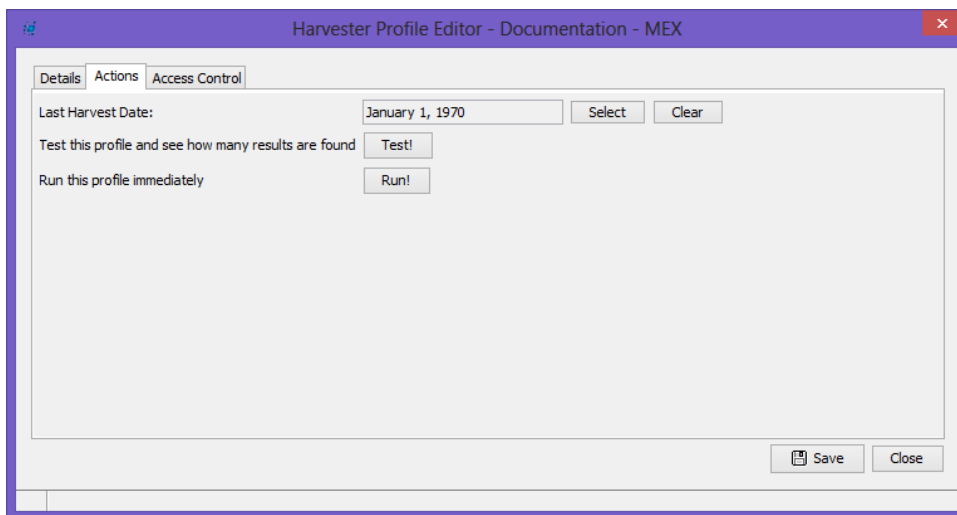






Figure 12 Harvester Profile Editor—Actions page

The Actions page elements include:

- **Last Harvest Date**—displays the date the harvester was previously run. Click  to choose a date and  to remove the date. The harvest will include all resources/metadata that have not been previously harvested from the repository for the time frame from the date displayed and the current date. The harvest can be scheduled to run at any time, refer to the [Harvester profiles scheduled task](#) section on page 17 for more information.
- —this function will test the profile settings and return the number of results without actually harvesting them.
 - a. To perform a test, click . The mouse cursor will turn into a spinner while the query is being performed.

- b. Once finished, a dialog displays informing of the number of returned resources, as shown in Figure 13. If this dialog does not appear, the harvester profile settings are configured incorrectly.

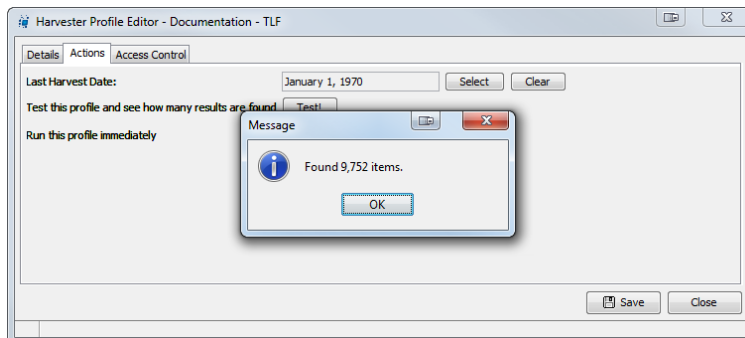


Figure 13 Successful harvest test run

- **Run!**—select this button to immediately run the profile. Run the harvest after performing a successful test.
 - a. To perform a harvester, click **Run!**. A dialog will pop up asking for confirmation. See Figure 14.

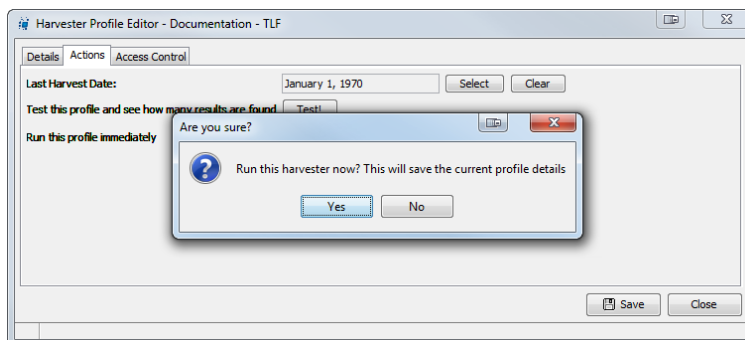


Figure 14 Run harvest confirmation dialog

- b. Once a harvest has been run, is not easily undone. Make sure you definitely want to run it before clicking **Yes**. A message dialog replaces the confirmation dialog, informing you that the harvester is now running in the background. See Figure 15.

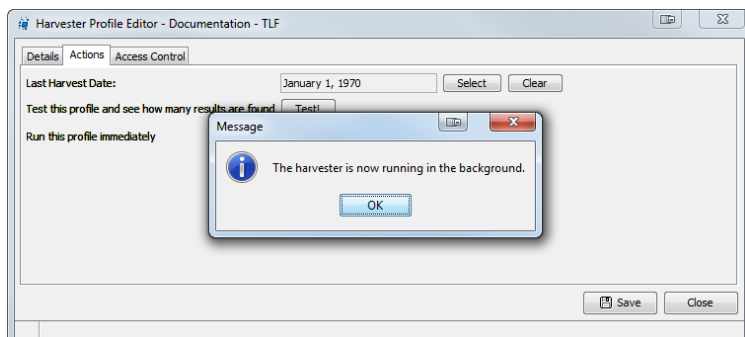


Figure 15 Successful initiation of harvest

- c. The harvester is now running in the background and will not require further user input. Harvester actions are logged in the resource-centre logs and can be viewed by accessing the file:

`C:\<path-to-equilla-install-directory>\logs\resource-centre\<date>\application.html`

- d. In this file harvester actions are recorded and can be viewed. A typical block of logged actions associated with a resource upload will look similar to that in Figure 16.

10:16:41.918	INFO	AbstractHarvesterProtocol	Uploading resource Meet a scientist: geologist and environmental scientist [no spoken instructions]
10:16:41.960	INFO	FileSystemServiceImpl	Creating combined manifest
10:16:41.979	INFO	FileSystemServiceImpl	Non-existent manifest requested for Staging/107/6994777b-14f5-0210-731e-e467213f4b88
10:16:42.242	INFO	FileSystemServiceImpl	commitFiles: from c:\equella\install\5\filestore\Staging\107\6994777b-14f5-0210-731e-e467213f4b88
10:16:42.242	INFO	FileSystemServiceImpl	commitFiles: to c:\equella\install\5\filestore\Institutions\documentation\Attachments\45\8b3c31e-dba-05a0-a63a-040d0443c0481
10:16:42.545	INFO	BackgroundIndexerImpl	Indexing batch of 1 items
10:16:42.742	INFO	ItemService	Finished operateAll on: class com.tle.core.workflow.filters.ArchiveOldFilter
10:16:42.742	INFO	AbstractHarvesterProtocol	Upload complete

Figure 16 Typical block of logged harvester actions

- e. As soon as the harvester is initiated, resources are being added to the selected collection. Resources are uploaded one at a time and can be discovered within EQUELLA before the harvest has fully completed.

Access Control tab

The **Access Control** page allows a harvester profile to be restricted to a set of users, groups or roles.

On this page, two privileges can be configured:

- **DELETE_HARVESTER_PROFILE**—delete this harvester profile.
- **EDIT_HARVESTER_PROFILE**—edit this harvester profile.

Other privileges relevant to Harvester Profiles, which are configured in the Security Manager, not from this page, are:

- **DOWNLOAD_ITEM**—download resources into collection(s).
- **CREATE_HARVESTER_PROFILE**—create a harvester profile.

Typically these privileges are granted to Administrators, but anyone who performs a harvest will need **DOWNLOAD_ITEM** granted to them for the given collection.

Further information on managing privileges is available in the *EQUELLA Security Administration Guide*.

To configure a privilege

1. Select the **Access Control** tab to display the **Access Control** page.
2. Click the **Who can?** drop-down arrow to reveal the list of available privileges.
3. Select the **delete this harvester profile** privilege.
4. Select the **Advanced** option to display the access control list.

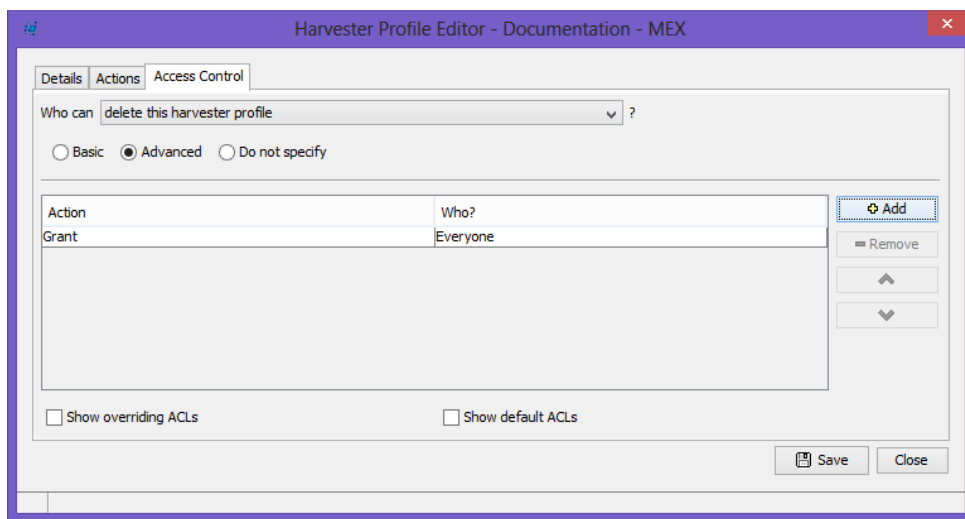
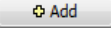





Figure 17 The Access Control page

5. Click .
6. In the **Action** field, select either **Grant** or **Revoke** by clicking on the current setting and selecting the appropriate option from the drop-down menu.
7. In the **Who?** field, click **Everyone** to display the **Select Recipients...** dialog. An example is shown in Figure 18.
8. Select the **Roles** option.
9. Click  to display a list of all available roles. Alternatively, type the first letter or letters of search resource and click .
10. Select **System Administrator** then select the  button to move the role to the right-hand pane, as shown in Figure 18.

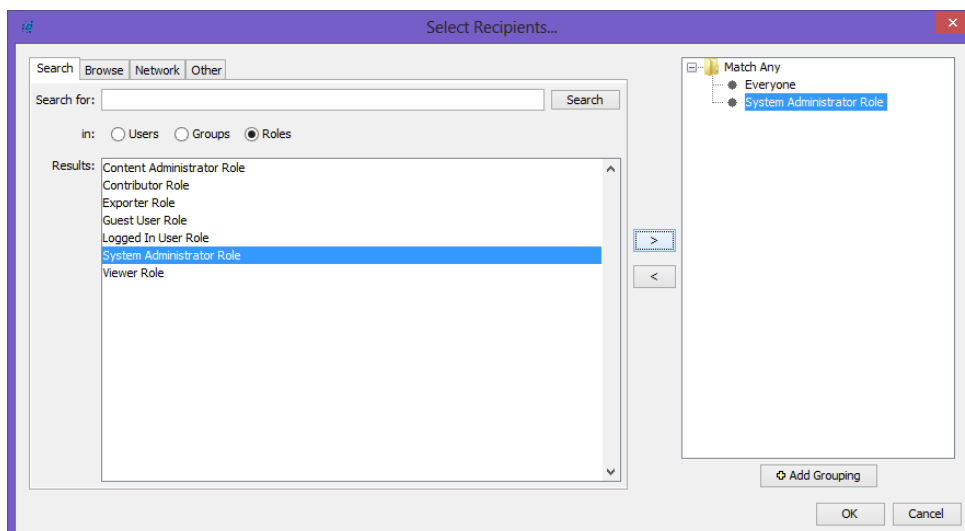





Figure 18 Select Recipients... dialog

11. Remove **Everyone** from the right-hand pane by selecting it then clicking the  button.

12. Click  to close the dialog. The delete privilege has now been granted to users in the system administrator role, as shown in Figure 19.
13. Click  to save the harvester profile.

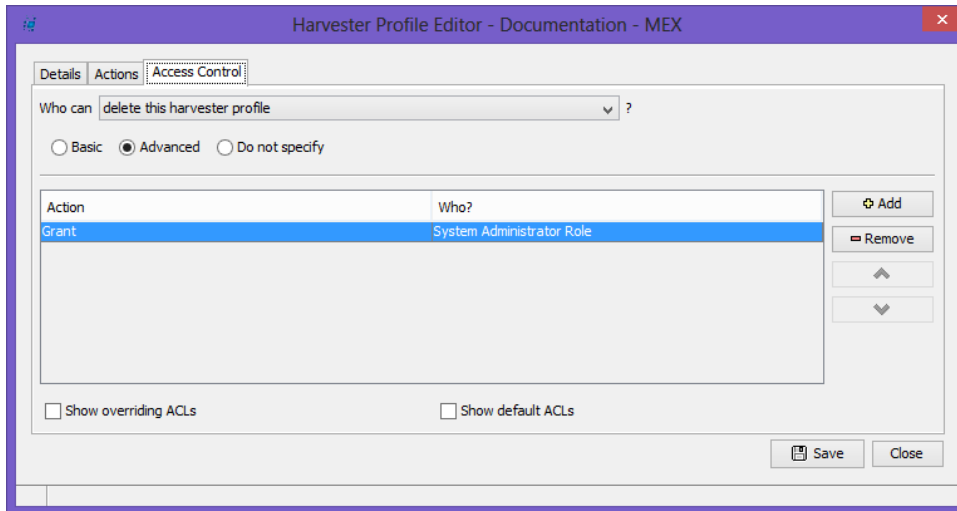
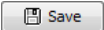



Figure 19 Completed access control configuration

Further privileges and users can be configured using the above procedure. Alternatively, refer to the *EQUELLA Security Guide* for more information on security configuration.

To save a profile

When all elements have been configured for a profile:

1. Click  to save the harvester profile details.
2. Click  to close the Harvester Profile Editor.

Harvester profiles scheduled task

Harvesters run with other EQUELLA schedules tasks. **Scheduled tasks** are configured from the **Settings** page.

To configure Scheduled tasks

1. Select **Settings** from the navigation menu, then click the **Scheduled tasks** link, as shown in Figure 20.

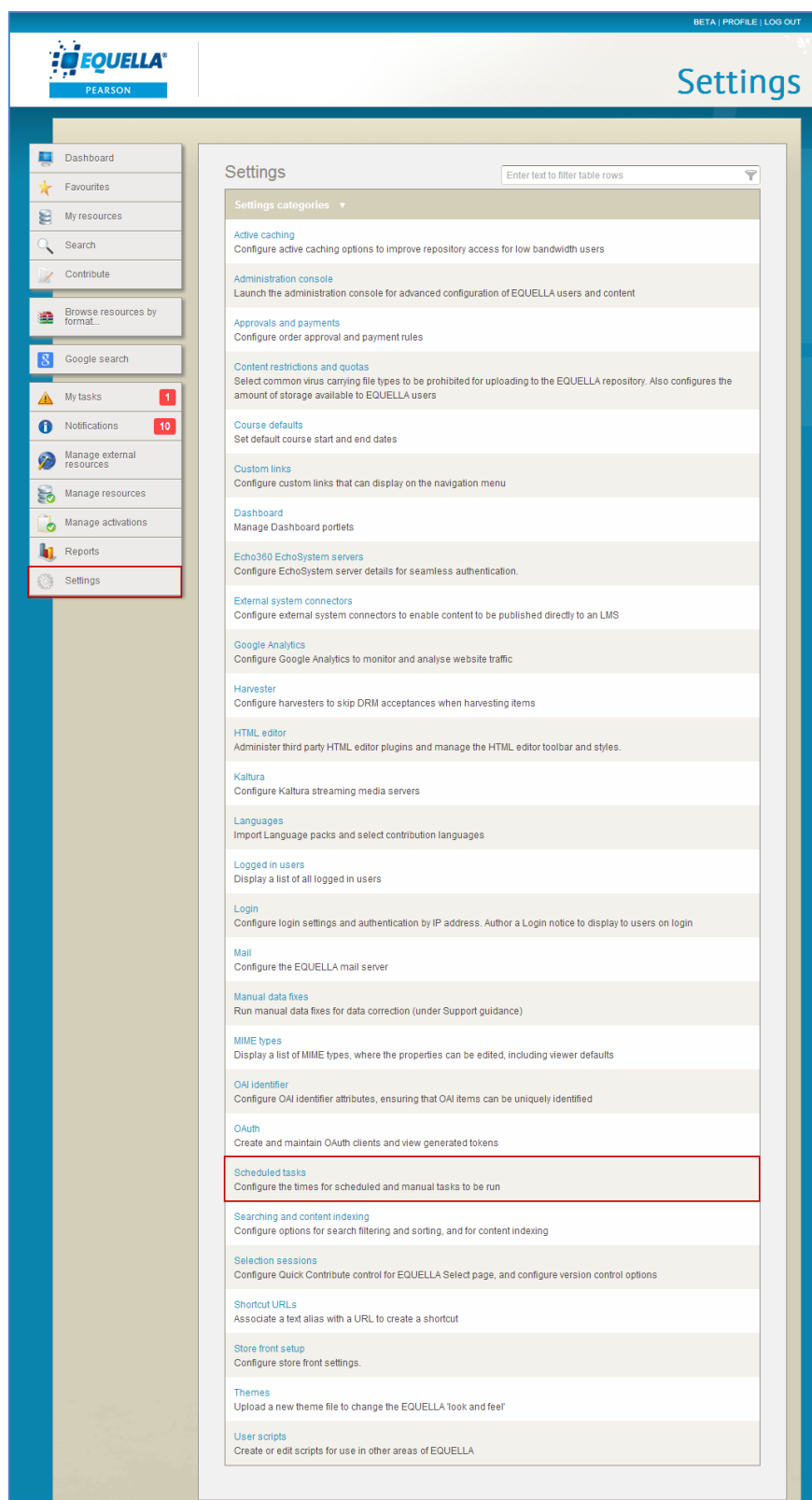


Figure 20 Settings page

The **Scheduled tasks page** displays, as shown in Figure 21.

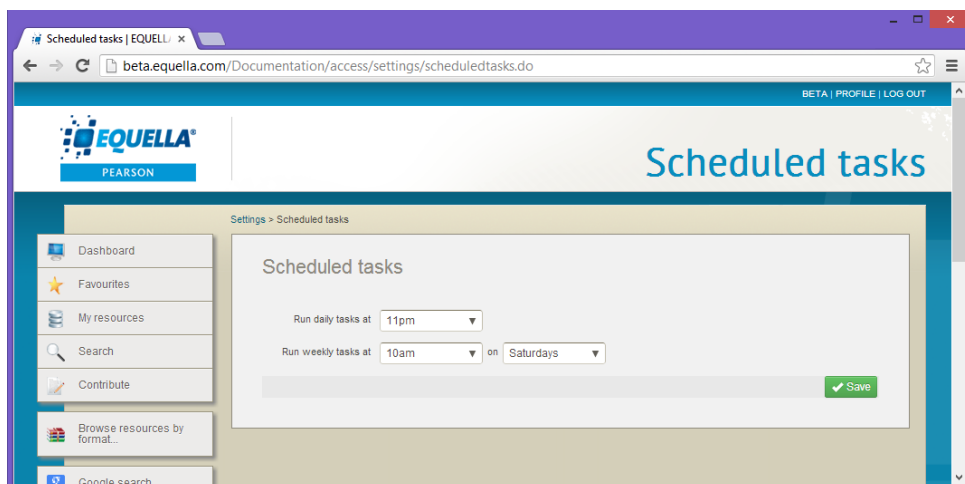



Figure 21 Scheduled tasks page

2. Harvester profiles are run daily. Select the required time from the **Run daily tasks at** drop-down, then click .

All enabled profiles are included when the scheduled task is run.

Further information on scheduled tasks is provided in the *EQUELLA Administration Overview Guide*.

Harvester system settings

A **Harvester** setting is available from the **Settings** page, and allows harvesters sourcing content from the current institution to harvest content without accepting any DRM licences.

To configure the Harvester option

1. Select **Harvester** from the **Settings** page.
2. Select the **Harvester** link from the Settings page. The **Harvester** page displays, as shown in Figure 22.

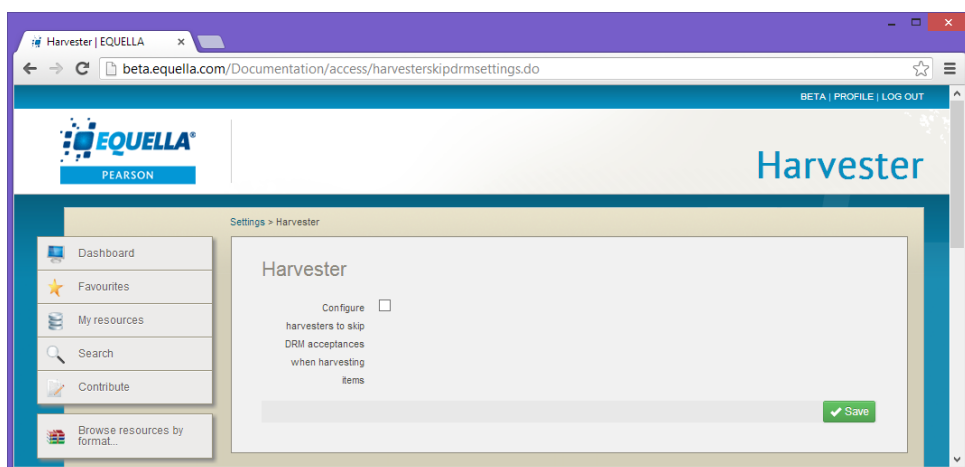


Figure 22 Harvester page

3. Select the **Configure harvesters to skip DRM acceptances when harvesting items** checkbox.
4. Click .

Contact Client Support

We are always happy to help.

If your organisation has a support agreement with EQUELLA then help is available at <http://equella.custhelp.com>.