Submitting a dataset to ANU Data Commons

RESPONSIBLE AREA: University Librarian, ANU Library
CONTACT: repository.admin@anu.edu.au
UPDATED: 15 October 2018

Contents

Data Commons Data Model

Data Commons is designed primarily to store data about research projects, research datasets, services and researchers. The system stores this information as a **Record**. Each record can be of one of the following types:

- **Activity:** An activity record contains information about a research activity that generates one or more outputs.
- **Collection:** A collection record contains information about one or more datasets or other related digital objects.
- **Service:** A service record contains information about a service that assists or supports a research activity.
- **Party:** A party record contains information about a person or an organisation that performs a role in relation to a research activity, collection or service.

A collection record is different from other record types as it allows digital objects in the form of data files that it describes to be uploaded to the system for storage. For data already hosted on other systems, a reference can be stored as one or more URLs that point to the network locations of those datasets or files.

Data Commons Security Model

Data Commons uses a security model to ensure that only authenticated users are able to perform tasks limited by the permissions assigned to them. Each user in the system can be assigned 0, 1 or more groups. For each group that is assigned to the user a set of permissions is assigned to determine the level of access that the user has to the records that belong to that group. Groups also follow a tree-like hierarchy where a user's permissions for a specific group allows the same level of access to all child-groups of that group, but not its parent group. A user who has permissions for no groups will only be able to see published records even when logged in.

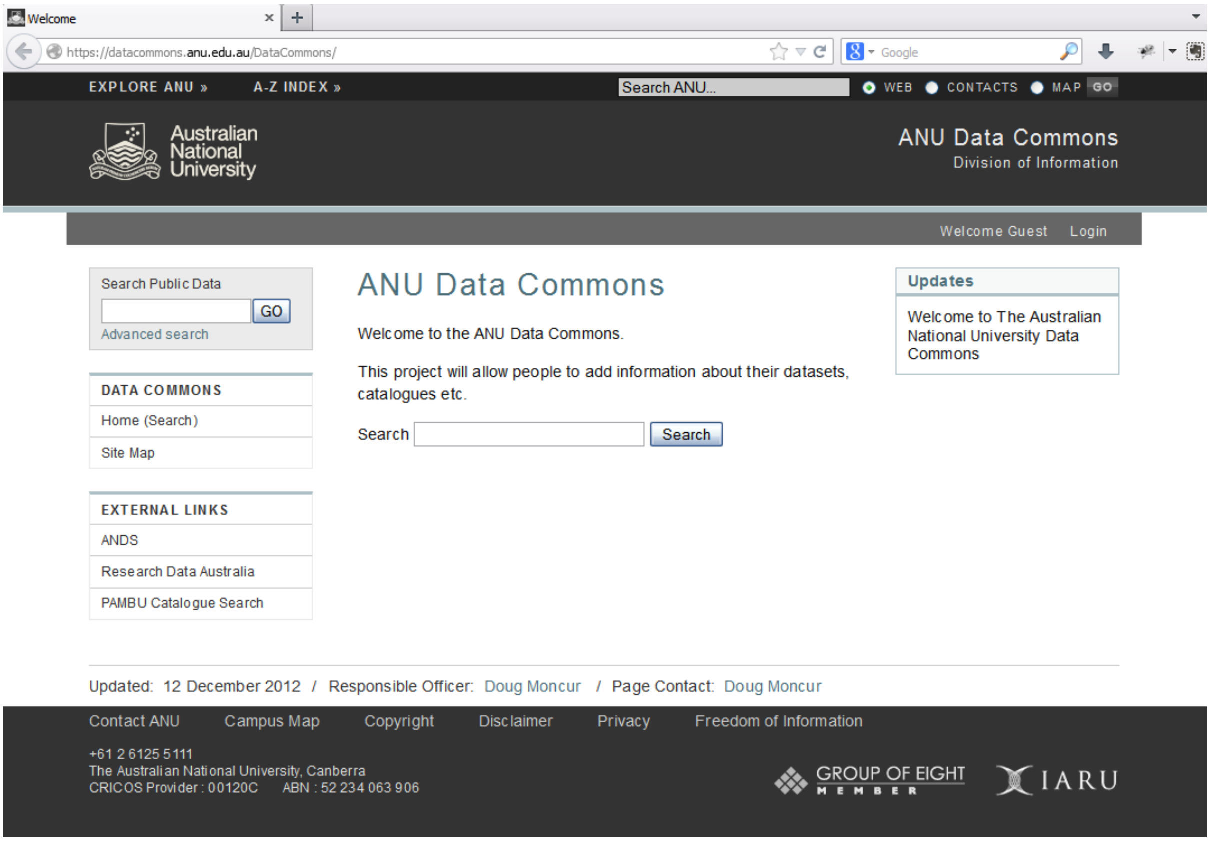
When a record is created, a group is assigned to it as its owner. All members of that group will then be able to access the record limited by their own permissions set for that group.

Features at a Glance

- Ability to store details of research activities, resources, and people as records in the system.
- Ability to store datasets and collection data in the system for digital preservation of data.
- Ability to restrict the accessibility of records.
- Workflow to allow users to request access to datasets and allowing access to requestors on a case by case basis.
- Ability to share records with other systems, such as Research Data Australia.

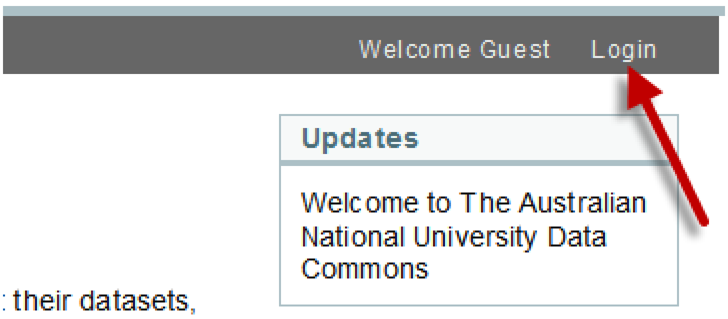
Accessing the Web Site

Type https://datacommons.anu.edu.au/DataCommons in the address bar of your web browser and press Enter. The home page similar to the one below will appear:



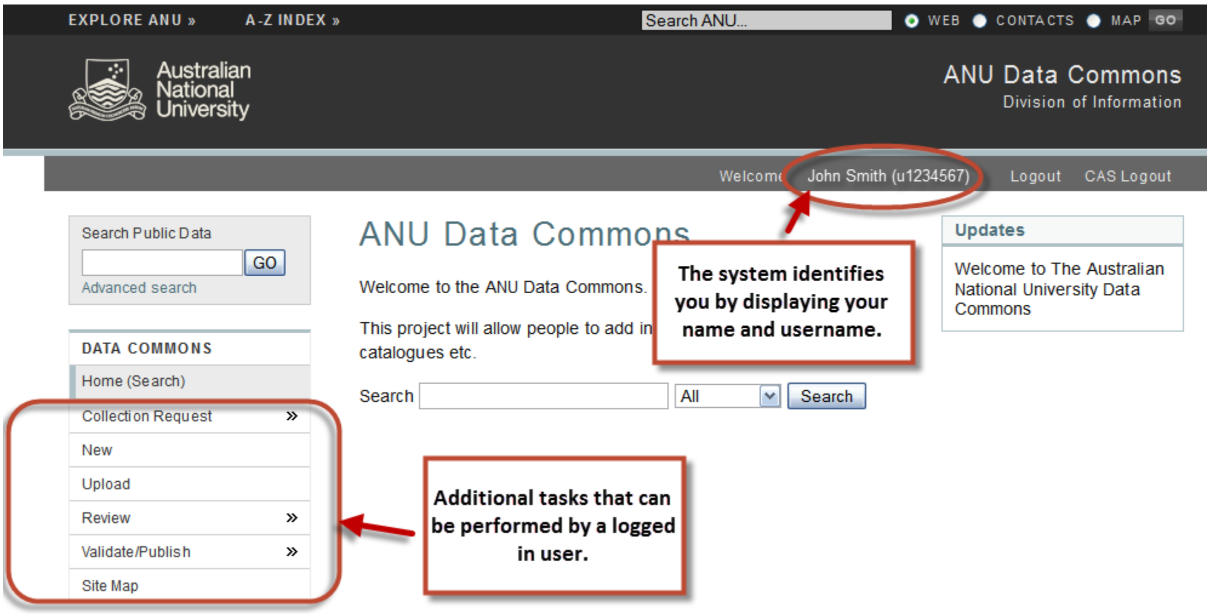
Logging in

To log into the system you either must have a valid ANU university ID and password, or must be registered as a registered user with an external email address as username. Click on the Login link in the top right corner. Contact the system administrator to discuss your access requirements.



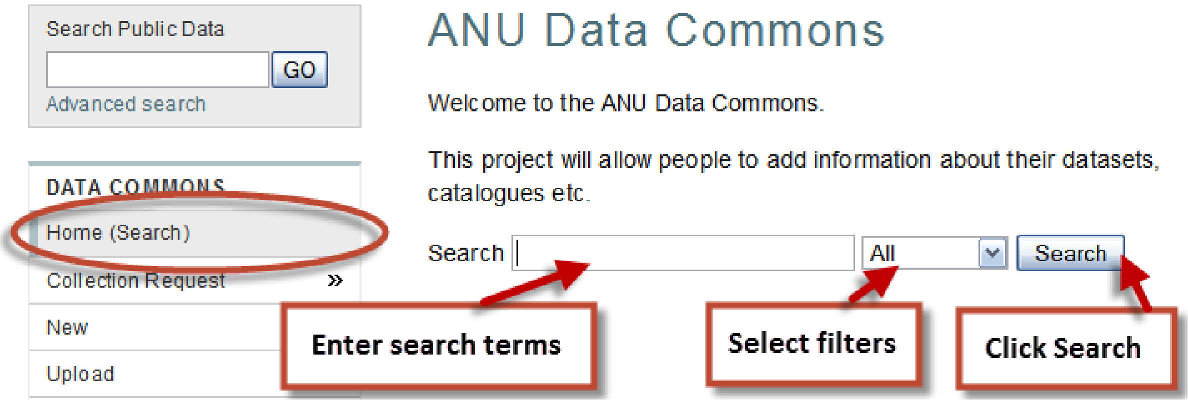
Once logged in, the home page should identify you by your display name and your username in the login bar. The menu area may list additional tasks that can be performed depending on the permissions assigned to the user in the system.

Contact repository.admin@anu.edu.au to request a change to your permissions, if required.



Search

Once logged in, a user can search all records that belong to groups to which the user has access. To search for records, click on the Home (Search) link in the menu area on the left of the screen that will open a page similar to the following:



Enter search terms in the search textbox and select a filter, if required from the dropdown list. One of the following filters can be selected:

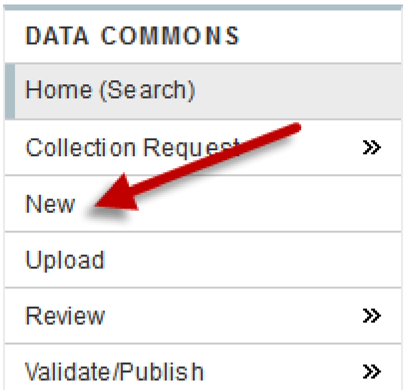
- **All:** includes all records that the user has permissions to.
- **Published:** restricts search to those records that have been published.
- **Team:** restricts search to those records that belong to the user's team.

Then click Search. The search results appear as below listing the title and brief description of all records that the user has access to, meet the filter criteria, and contains one or more of the search terms specified:

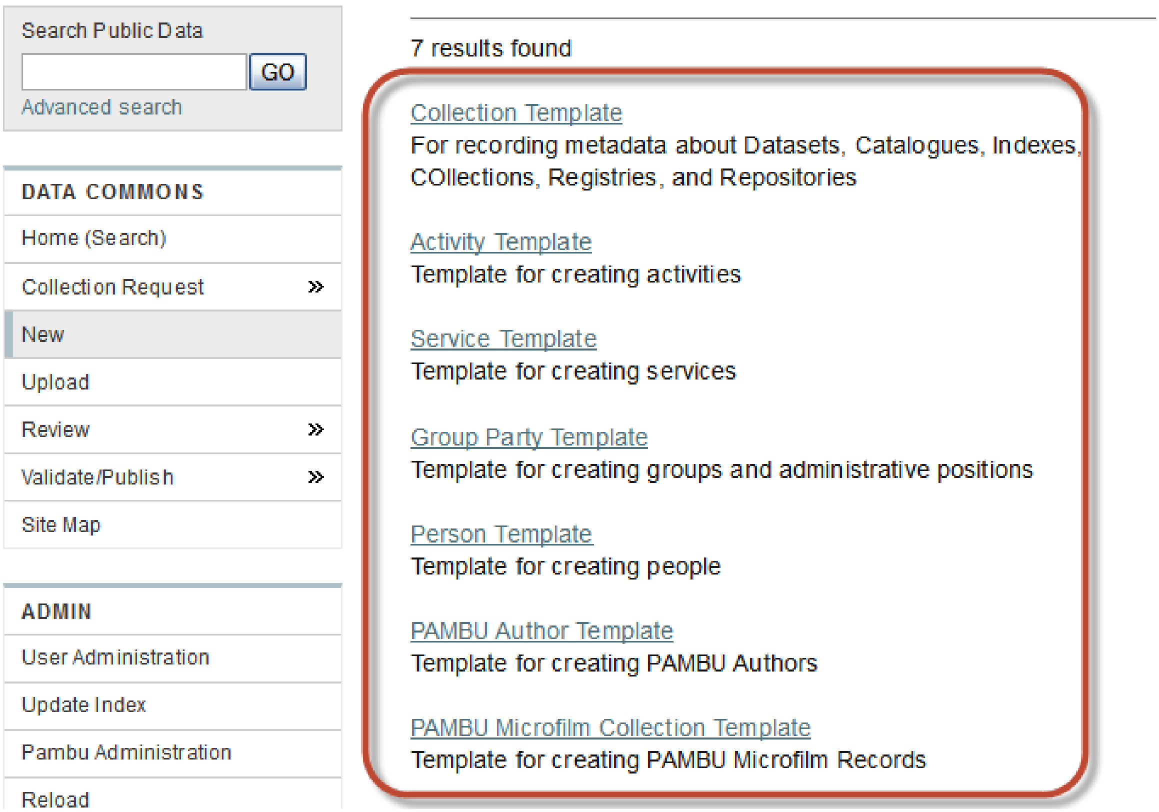
Records

Add a New Record

To add a new record into Data Commons, click on the **New** menu item.

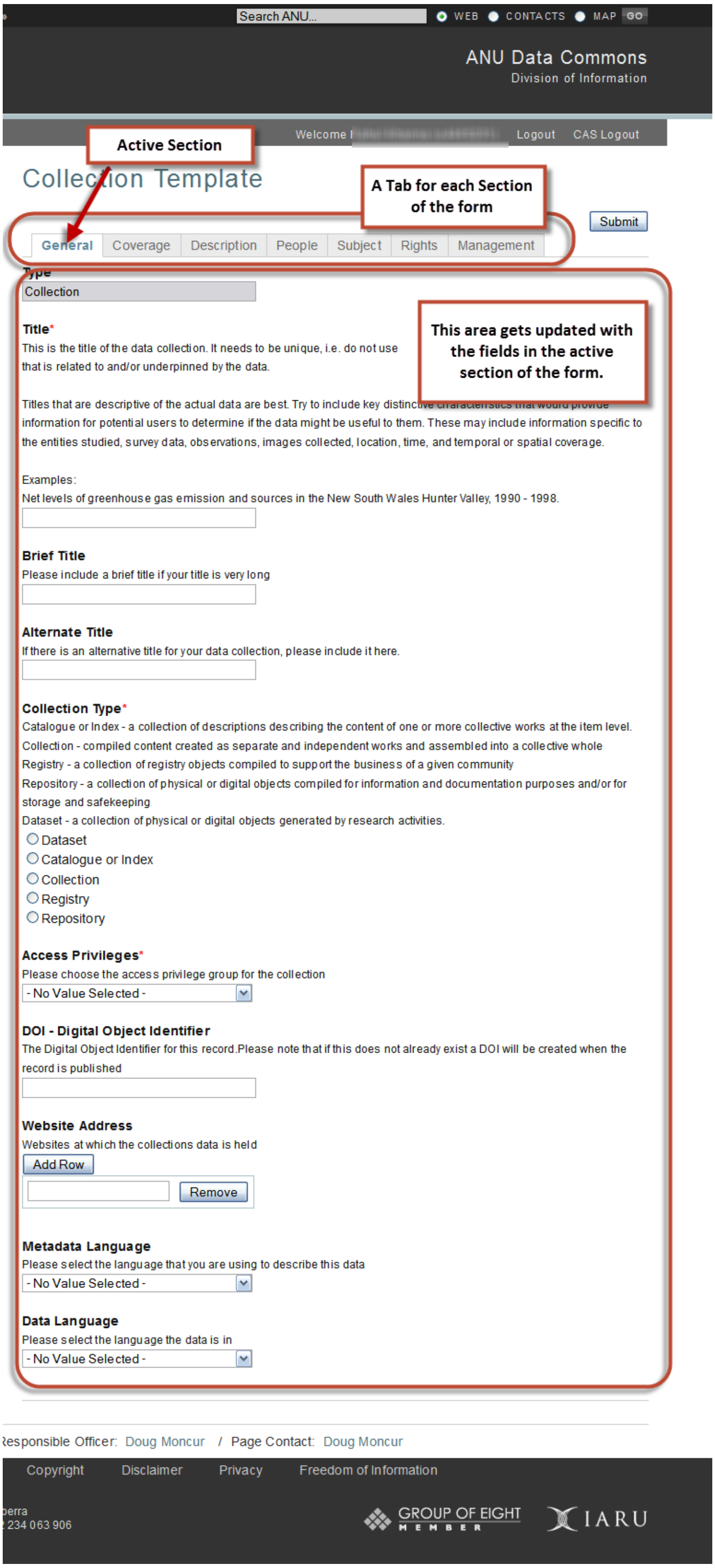


Click on the template corresponding to the type of record to be created.



***Note:** Created records can be linked with each other and do not require creation in a specific order.*

The next page lists a form that must be filled out for a record to be created. Some fields will be prefilled, some are mandatory, while the rest are recommended, but optional. The form will be divided into sections. Each section can be accessed by clicking on its relevant tab on top.


Search ANU...
WEB
CONTACTS
MAP
GO
ANU Data Commons
Division of Information
Welcome
Logout
CAS Logout
Active Section
Collection Template
A Tab for each Section of the form
Submit
General
Coverage
Description
People
Subject
Rights
Management
Type
Collection
This area gets updated with the fields in the active section of the form.
Title*
This is the title of the data collection. It needs to be unique, i.e. do not use
that is related to and/or underpinned by the data.
Titles that are descriptive of the actual data are best. Try to include key distinc
information for potential users to determine if the data might be useful to them. These may include information specific to
the entities studied, survey data, observations, images collected, location, time, and temporal or spatial coverage.
Examples:
Net levels of greenhouse gas emission and sources in the New South Wales Hunter Valley, 1990 - 1998.
Brief Title
Please include a brief title if your title is very long
Alternate Title
If there is an alternative title for your data collection, please include it here.
Collection Type*
Catalogue or Index - a collection of descriptions describing the content of one or more collective works at the item level.
Collection - compiled content created as separate and independent works and assembled into a collective whole
Registry - a collection of registry objects compiled to support the business of a given community
Repository - a collection of physical or digital objects compiled for information and documentation purposes and/or for
storage and safekeeping
Dataset - a collection of physical or digital objects generated by research activities.
Dataset
Catalogue or Index
Collection
Registry
Repository
Access Privileges*
Please choose the access privilege group for the collection
- No Value Selected -
DOI - Digital Object Identifier
The Digital Object Identifier for this record.Please note that if this does not already exist a DOI will be created when the
record is published
Website Address
Websites at which the collections data is held
Add Row
Remove
Metadata Language
Please select the language that you are using to describe this data
- No Value Selected -
Data Language
Please select the language the data is in
- No Value Selected -
Responsible Officer: Doug Moncur / Page Contact: Doug Moncur
Copyright
Disclaimer
Privacy
Freedom of Information
GROUP OF EIGHT
MEMBER
IARU

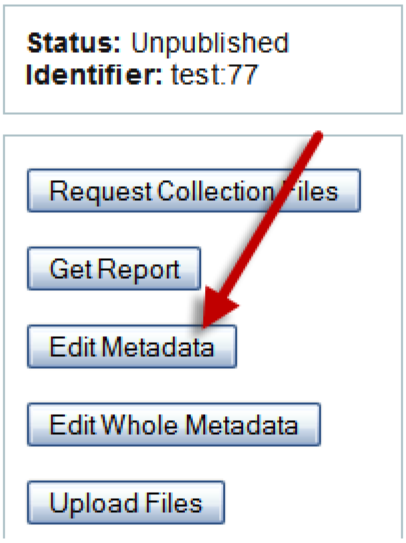
Note: It is highly recommended that all relevant fields be filled out to describe the record as verbosely and accurately as possible. Providing more details about a record makes it easier to find.

Once the form has been filled out, click on the Submit button in the top right of the window, which will create the record with the specified data and take you to that record's page that looks like the following:

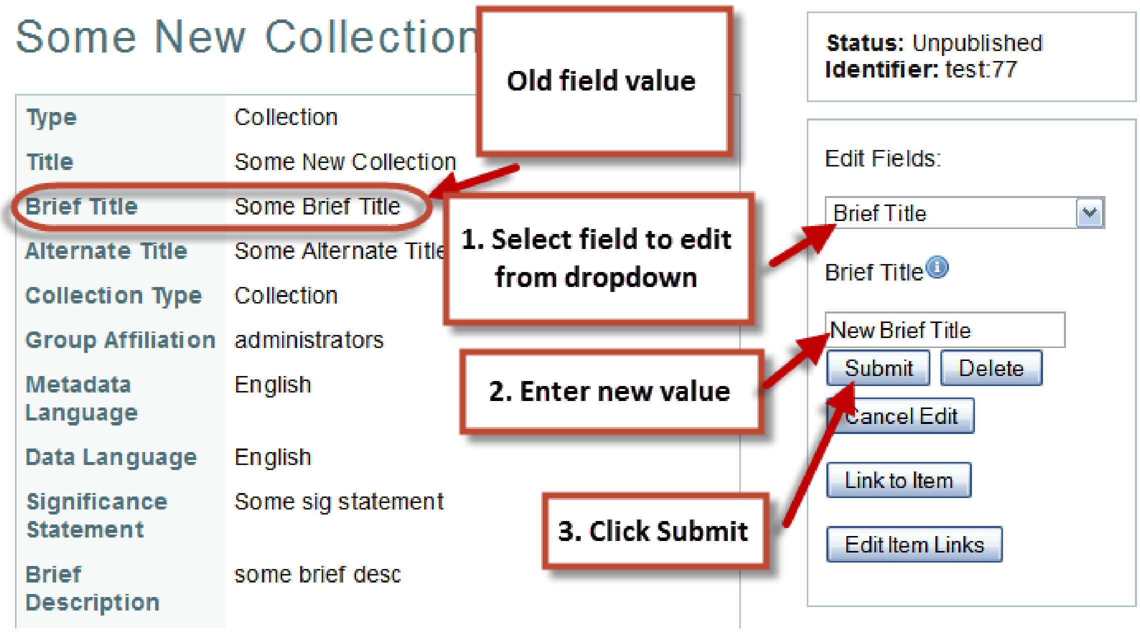
Note: *Any link can be opened in a new tab by holding the Control key down while clicking on the link. To open a link in a new window hold down the Shift key while clicking on the link.*

Edit a Single Field

To edit the value of a single field, click on the Edit Metadata button.

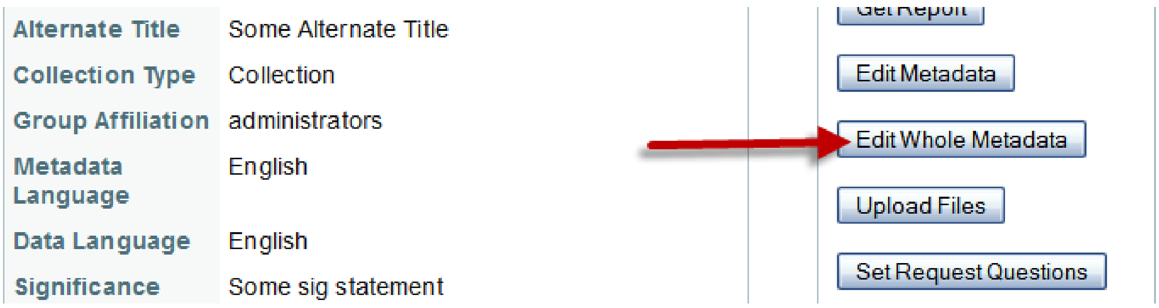


Then from the Edit Fields dropdown select the field whose value is to be changed, select or enter the new value in the area below it, then click the **Submit** button.



Edit Multiple Fields

If you want to update multiple fields in a record, editing individual fields as explained in section Edit a Single Field, can be quite cumbersome. To edit multiple fields at one time, click on **Edit Whole Metadata** button on the record's page. This will open the record's form just as when the record was being entered for the first time as explained in section Add a New Record.

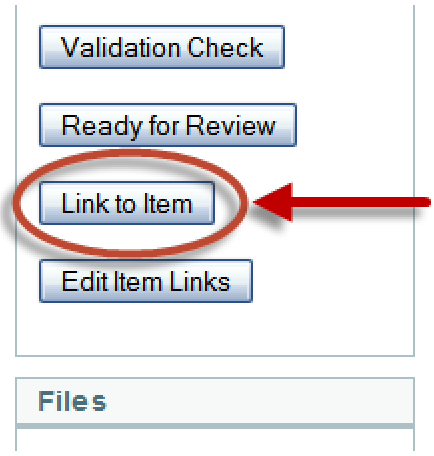


Make the changes to one or more fields as required, then click Submit button in the top right corner.

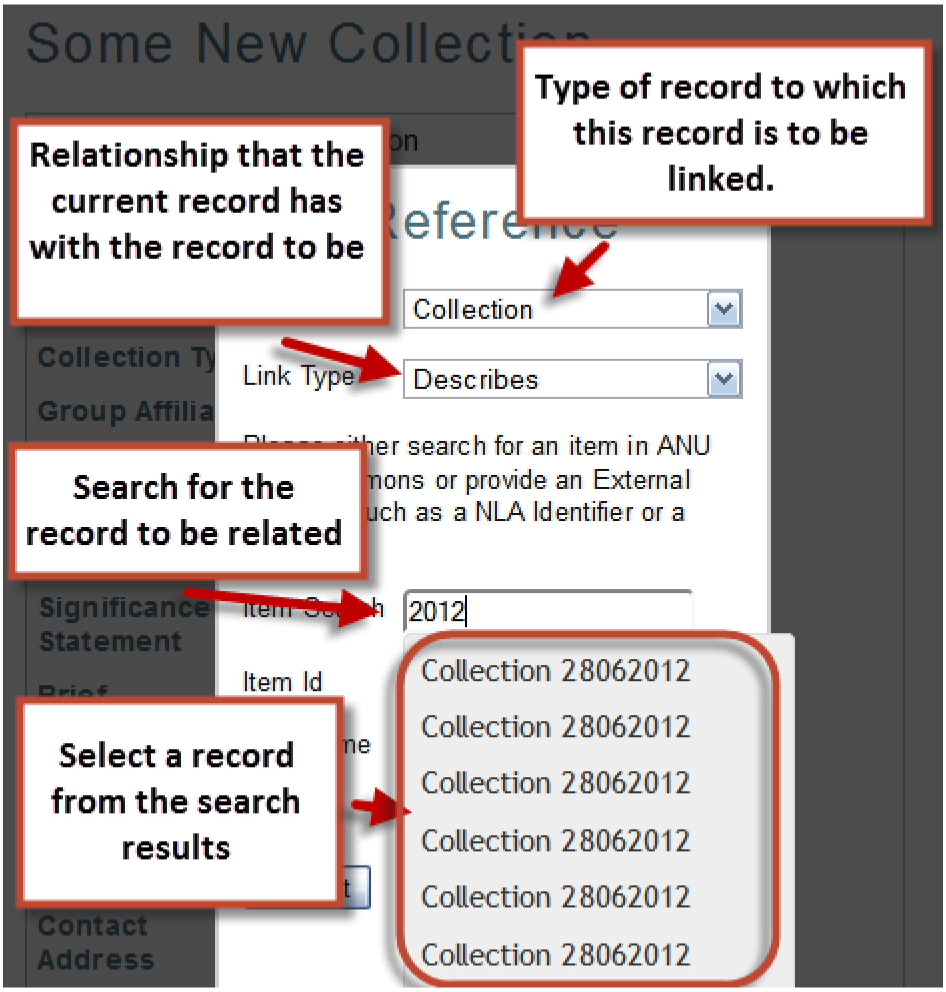
Linking Records

A relationship can be established between two records that have something in common. The commonality is described by the relationship itself. For example, a collection may belong to a research project that is managed by a researcher. Linking such records with each other enables discovery of all related records.

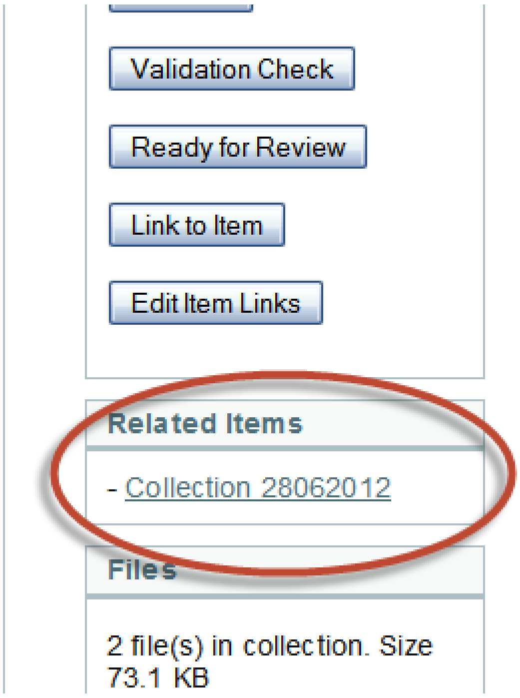
To link a record with another record, click on the **Link to Item** button on the record's page.



That pops up a dialog similar to the one below:



Click on the type of the target record – Activity, Collection, Party or Service. Then select the link type that describes the relationship between the **current record** and the **target record**. The available options depend on the type of current record and the target record. Then enter a few search terms to search for the target record. Once a list of search results appear, select one of them as the target record. That automatically populates the Item ID and Item Name fields. Click the **Submit** button to save. Once saved, all linked records show up on the record page as follows.

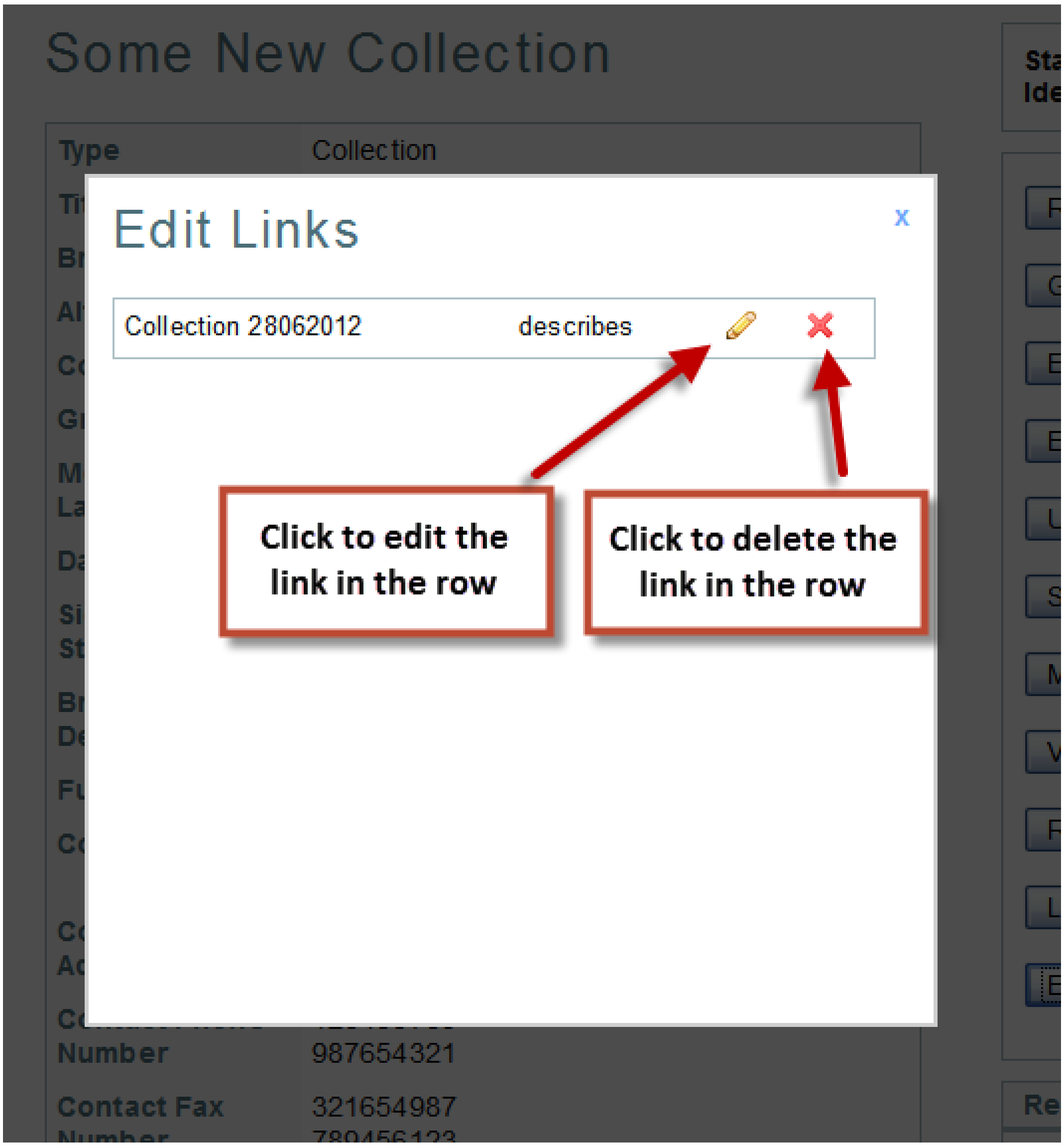


You can click on the link to navigate to a linked record's page. Repeat the steps above to link additional records.

> *Linking a record to another record creates a bidirectional link. For example, when collection A describes Collection B and a link is created to represent the relation between A & B, a corresponding link is created that links that translates to Collection B is described by Collection A.*

Modify Links

To modify existing links between records, navigate to either of the two linked record's pages, then click on **Edit Item Links** button. A dialog appears similar to the following that lists all links to other records.



To make changes to a link click on the edit icon as shown above which will display the item link dialog box. Follow the steps in the section Linking Records to make changes as required. Then click on **Submit** button to save changes.

To delete a link, click on the delete icon beside the link to be deleted as shown above.

Data Files

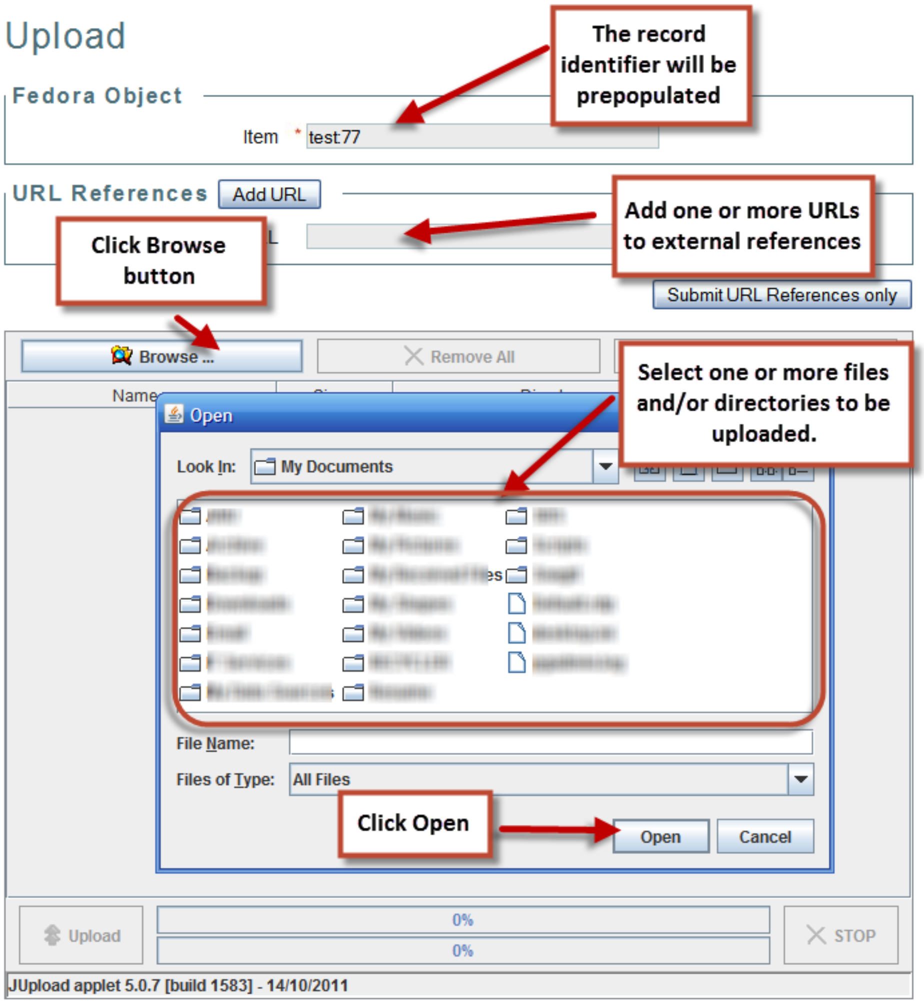
Upload files to Collection

The system allows you to upload files to records of the type Collections. This allows storage and preservation of digital objects that a collection comprises of. Data can be either stored in the system itself or a reference can be stored in the form of a URL that resolves to the location of the data.

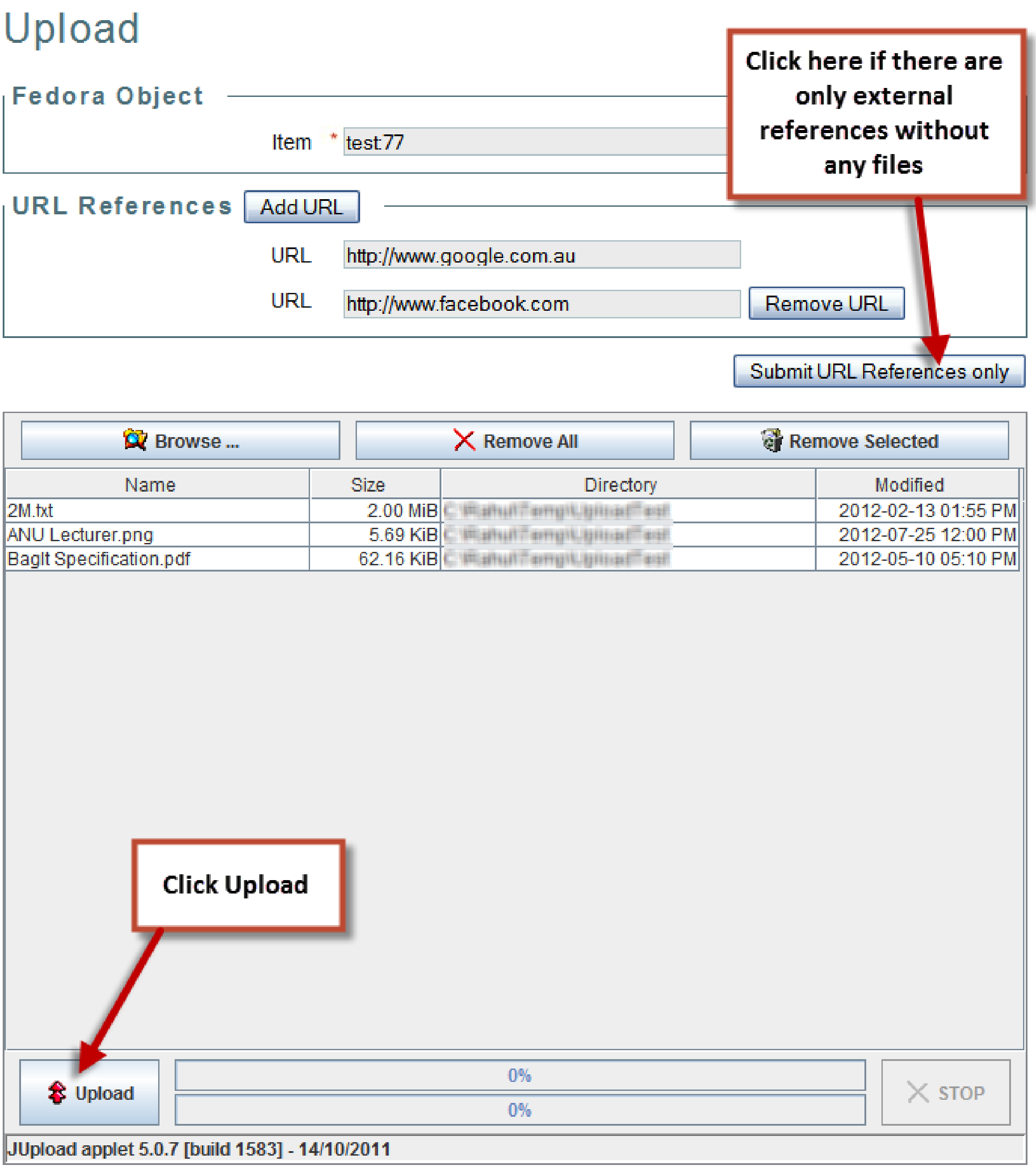
To upload one or more file to a collection, click on the Upload Files button on the record's page.

On the following page, enter one or more URLs to external references, if any. For each additional external reference, click on the Add URL button to add a new row to enter an external URL. Leave the textbox blank, no external references are to be added. Then click on the browse button to select files and directories to upload.

> *Note: The upload page uses JUpload Java Applet to upload files to the server. The web browser must be configured to run Java Applets. In most cases, simply installing the latest version of the Java Runtime Environment from http://www.oracle.com/technetwork/java/javase/downloads/index.html should suffice as the Java installer automatically installs the necessary plugins in web browsers installed on the system. In other cases, refer to your web browser's user guide to get the plugin to work with the browser. Some browsers actively prevent you from using an older version of Java. If that's the case, download and install the latest version of Java to re-enable the plugin.*



Then click on Upload.

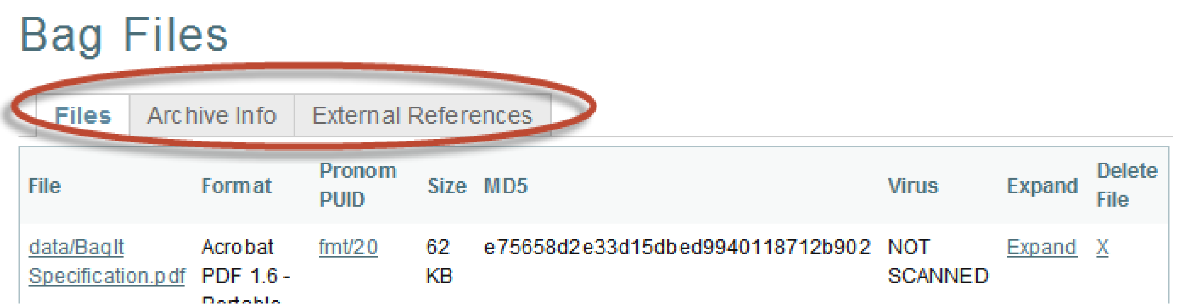


If there aren't any files to upload but only external references, then click on the button **Submit URL References only**.

> *This process only adds files to a collection. Existing files in a collection remain untouched unless one or more files have the same name as the files already uploaded. If newly uploaded file has the same name as an existing file in a collection, the existing file is archived and replaced with the uploaded file.*

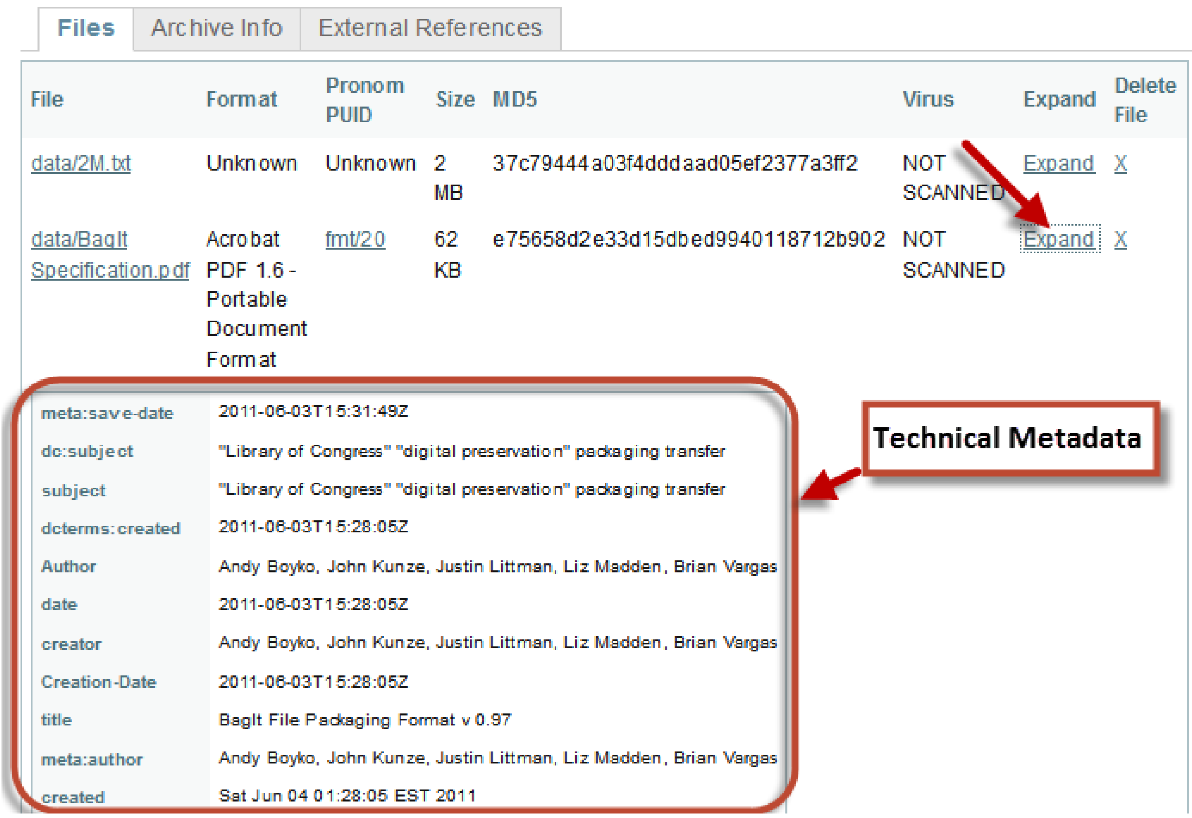
During the upload process, the progress bar at the bottom continually updates to show upload progress. Once the upload has finished, the Bag Files pages opens that lists the files and external references that have been uploaded. The bag page that that time only lists the files that have been uploaded. Additional information such as MD5 checksum, virus scan and metadata extraction happens as a separate process in the background and may take a while before their results appear on the Bag Files page.

Information about the files uploaded to a collection is categorized into sections that can be accessed by clicking on the relevant tab.



The **Files** tab lists information about each file stored in the system against the record. A file's name, format, Pronom Format ID, size, MD5 checksum, Virus Status and technical metadata can be viewed on this page.

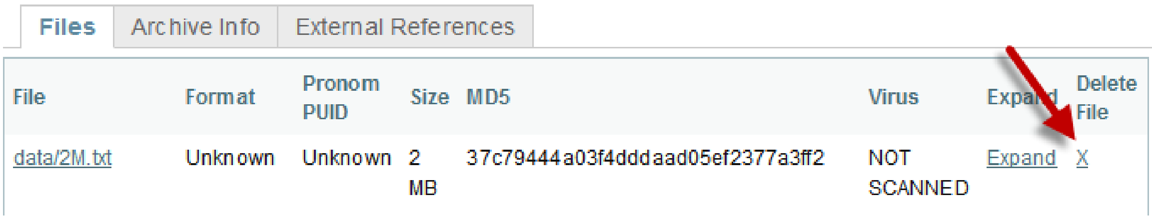
Technical Metadata of a file is hidden by default. To view this information about a file, click on **Expand** for the file whose information you'd like to view.



The system's ability to recognise a file format depends on that format's presence in the PRONOM file identification database. Similarly, the system's ability to extract *metadata from a file depends on Apache Tika's support for the file format. Refer to http://tika.apache.org/ for more information.*

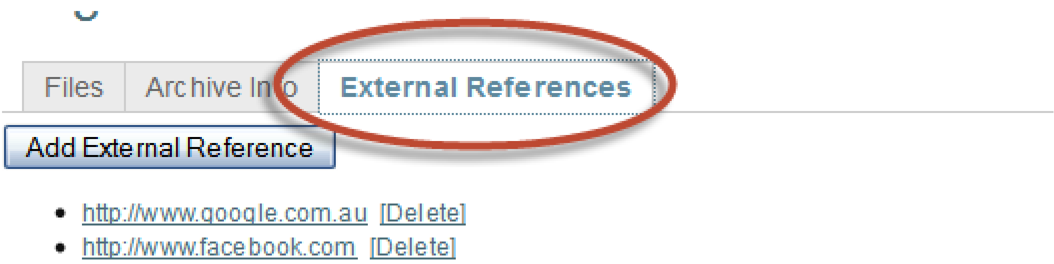
Deleting files from a Collection

To delete a previously uploaded file in a collection, navigate to its Bag Files page. Then click on the **X** link against the file to be deleted. Click OK in the confirmation dialog box. Once the file's deleted, the page reloads displaying the files that remain in the collection. These steps can be repeated for each additional file to be deleted.



Managing External References

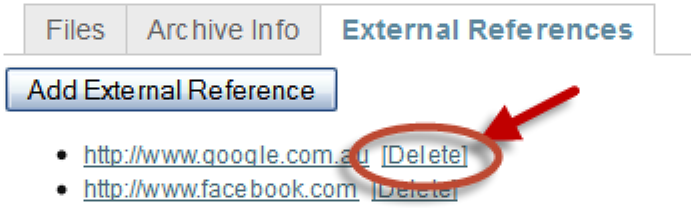
In addition to being able to store files that belong to a collection, the system provides the means to refer to files stored outside the system. This reference is in the form of a URL that points to a network location of said file. External references can be viewed by clicking on the External References tab on the Bag Files page.



To add an external reference, click on the button Add External Reference, enter a URL in the dialog box that follows, and click OK.

The page then refreshes displaying all external references associated with the collection.

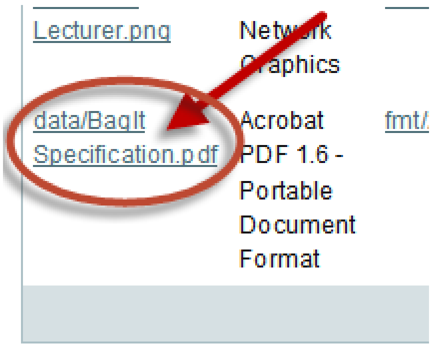
To delete an external reference, click on the **[Delete]** link beside the external link to be deleted. Click OK in the confirmation dialog that follows. The page will then refresh with the remaining list of external references.



Downloading Files from a Collection

Files uploaded to a collection can be downloaded individually or as a single ZIP file containing all files belonging to a collection.

To download individual files, open the Bag Files page and click on the link of the filename to be downloaded.



Depending on the configuration of your web browser, you will be prompted to open or save the file, or the file download will begin saving the file into the default download directory. Refer to your web browser's user manual for details.

To download all files in a collection as a single ZIP file, click on the **Download all as Zip** link at the bottom of the table that lists all files in the collection.

Depending on the configuration of your web browser, you will be prompted to open or save the file or the file download will begin saving the file into the default download directory. Refer to your web browser's user manual for details.

.