

Archives Hub


About Coverage Features Researching **Contributing** Search Help

- Contributing
- Data Creation
- Data Enhancement
- Data Administration
- Admin interface (CIIM)**
- Revising descriptions
- Training

CIIM

The CIIM is the administrative interface for the Archives Hub. The CIIM allows contributors to view the descriptions they have on the Hub, to un-publish and re-publish and to upload new descriptions that are Hub valid EAD.

Access
<https://admin.archiveshub.jisc.ac.uk/ciim>



CIIM Browse interface

HubbuB, November 2019: The CIIM, the data and the pipelines

What is the CIIM?



The CIIM is the system that the Archives Hub uses to process and index its data. It was developed and is maintained by Knowledge Integration (K-Int). You can find out more about it on the K-Int site: <https://www.k-int.com/products/ciim/>

It provides a unique environment for our workflow, where K-Int set up a certain amount of processing and validation, but the Hub team also include their own validation routines. This means we can create a whole range of 'pipelines' to process data from different sources.

11 Global Transformations					
1	EAD	Tidy/Trimmed EAD from EAD	Global Organisation for Workflow Transformation	19-Jun-2017 15:25 19-Jun-2017 15:25	✓ Enable
2	EAD	Folksonomy/headers EAD from EAD	Global Organisation for Workflow Transformation	28-May-2018 18:24 28-May-2018 18:24	✓ Enable
3	EAD	Reorder/label EAD from EAD	Global Organisation for Workflow Transformation	07-Oct-2017 18:11 07-Oct-2017 18:11	✓ Enable
4	EAD	Tidy/Longwords EAD from EAD	Global Organisation for Workflow Transformation	28-May-2018 18:25 28-May-2018 18:25	✓ Enable
5	EAD	DeDup/All EAD from EAD	Global Organisation for Workflow Transformation	25-Jul-2018 12:45 25-Jul-2018 12:45	✓ Enable
6	EAD	Tidy/Source/And/Links EAD from EAD	Global Organisation for Workflow Transformation	17-Apr-2018 18:28 17-Apr-2018 18:28	✓ Enable
7	EAD	Tidy/Group/All EAD from EAD	Global Organisation for Workflow Transformation	20-May-2018 18:21 20-May-2018 18:21	✓ Enable

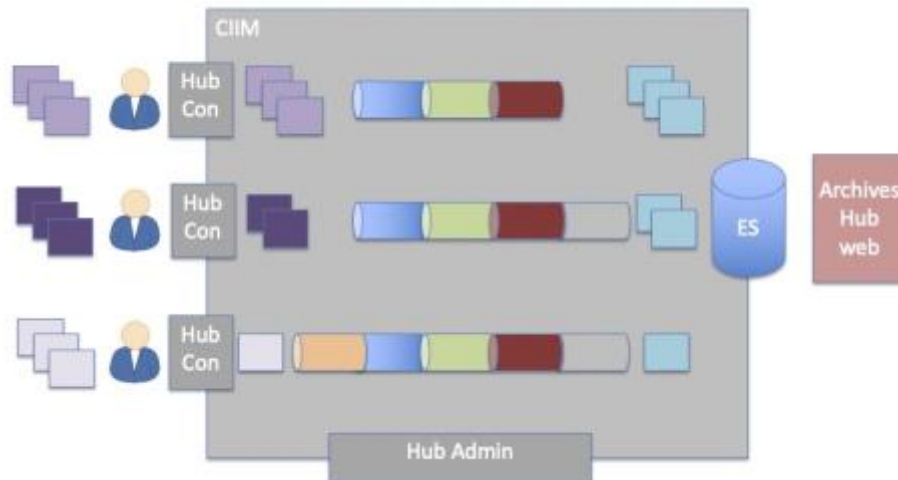
1 - Part of our global pipeline, showing a range of different processing routines

The CIIM has an entry for each contributor, which is created automatically when we add an 'EAG' file for them. We can give each contributor individual pipelines if necessary, though many contributors share the same set of pipelines.

The CIIM provides administrative access for the Hub team and access for each individual contributor, for them to see and work with their own descriptions. The CIIM can be used to browse, search, filter, un-publish, re-publish and delete descriptions.

Data is constantly being uploaded to the CIIM, and it is doing complex processing, so sometimes it does falter - it may appear frozen. If this is the case, try refreshing the page first of all. If the processing still does not complete, it is best to leave it for some time - maybe an hour - and then if it still seems stuck, email the Hub. It is likely we will already be aware of the problem and be addressing it.

All descriptions that are uploaded to the CIIM are backed-up, so we can always recover anything if need be.








2 - Workflow: contributors provides data that is processed through a set of transformation routines before it is made live.

Exporting from Archival Management Systems

Over the last 2 years we have spent a good deal of time creating pipelines to work with EAD exports. We have concentrated on Calm because it is used by many of our contributors, but we are working with other systems.

We have a set of pipelines that makes up a Calm group. All exports from Calm go through these pipelines when they are uploaded, and they are then validated before being made live.

⇅ Calm Transformations

1	EAD	TidyCalm tidy-calmead.xsl	 Calm Group
2	EAD	SplitListedVals split-listed-values.xsl	 Calm Group
3	EAD	MakeNormalisedDates make-normalised-date.xsl	 Calm Group
4	EAD	NormaliseUnitidLabel normalise-unitid-labels.xsl	 Calm Group
5	EAD	RepairCMRCodes repair-cc-mc-rc-attributes.xsl	 Calm Group

3 - Calm group with a number of transformation routines included.

For instance, you can see 'Make Normalised Dates' in the screenshot above. This is a small piece of code (XSLT) that interrogates the display dates and tries to create dates for searching. For example, if the display dates are 'January 1901 - March 1902' the normalised date that is created would be 190101/190203. Without a normalised date, the content will not be included in a date search.

Some of these transformations are complicated, and require careful analysis of the data, sophisticated XSLT and rigorous checking of the outcomes. One transformation may fix a problem, but if it is not done carefully and put in the right place, it may create a problem elsewhere.

The 'Tidy Calm' transform does a host of things (here things are getting a bit more technical):

- Selects unitid data based on unitidMap value
- Filters according to config settings (these settings can e.g. add dividers to subjects and names, and ensure 'closed' catalogues are not displayed)
- Cleans up eadid/@countrycode and eadid/@mainagencycode attributes
- Cleans up unitid/@countrycode and unitid/@repositorycode attributes
- Attempts to detect URIs in text and convert to extref (we often get descriptions with URIs that are not properly formed as hyperlinks)
- Fixes various invalid structures

We have been working on an export from AdLib, we are harvesting successfully from AtoM, and we are also looking at other systems, including Epexio and Soutron. These can all take time because what may work for one set of data from one repository does not work for the next set from a different

repository...or the next set, and so on. Archivists catalogue differently, sometimes within one repository, and data may have been migrated from old systems, and therefore have legacy issues.

Excel

We have tried for some time to work with Excel data. We have reached the conclusion (along with many others who have tried this) that we will have to provide a template. We cannot take in Excel data from different sources, as each one would require its own transformation, which would be very time-consuming. Additionally, creating an EAD hierarchy from a flat spreadsheet structure requires some kind of referencing to indicate a parent/child relationship. We can make sure this happens with our own Excel template.

Once we have done this, we will create pipelines for Excel data. It may take some time to complete this work, as we have to test behaviours for each field, think about multiple values in one field, think about things like hyperlinks and attribute values, and deal with the complexities of archival nesting.

Revisions and Developments

Pipelines do not necessarily stay unchanged. If we get a new contributor with different data, we may look at changing the pipelines to accommodate that data. The CIIM allows for constant change and flexibility. However, it is crucial that we continue with repeatable processing. So, if a contributor provides a revision to a description several years after the original description was uploaded, we have to deal with that, and make sure that versioning is sound.

In terms of CIIM developments, we are looking at work around names (name 'authorities') and also images (and IIF).

Access to the CIIM

All our contributors can have a login for the CIIM. Most contributors do have this now, and some who export their data from archival management systems, or use a native XML editor, are uploading descriptions. Contributors should only upload if they have agreed this with us, otherwise the correct pipelines may not be in place to ensure that descriptions are valid and publish OK.

If you use the EAD Editor to create and revise descriptions, as soon as you send these to the Hub they will be uploaded, and you do not need to log into the CIIM. However, it may be useful if you want the opportunity to temporarily un-publish descriptions, or just to check what is currently on the Hub.

Email us if you don't have access and would like a login.

CIIM Guide and Tutorials

See the CIIM page on the Hub for a link to the Guide and video tutorials:

<https://archiveshub.jisc.ac.uk/ciim/>

The videos are on our YouTube page, where there is a playlist with our current CIIM tutorials for Browse, the Data Manager, uploading descriptions and problems with uploads.

Do provide feedback on these - we are happy to develop our help guides in the best way possible to be of use to our contributors.