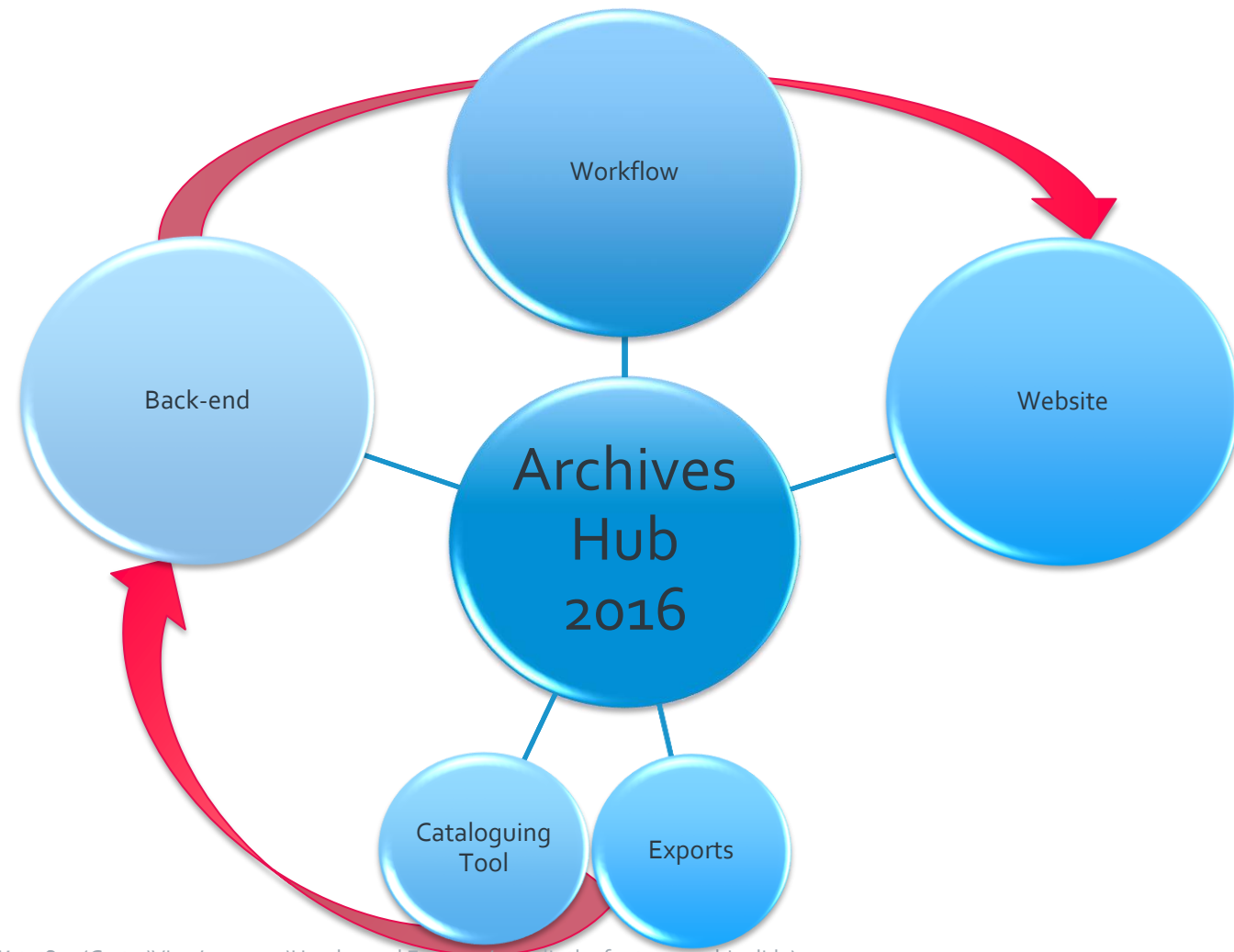


15/06/2018

# In Out, In Out, And Shake It All About

A Moving Story of Data



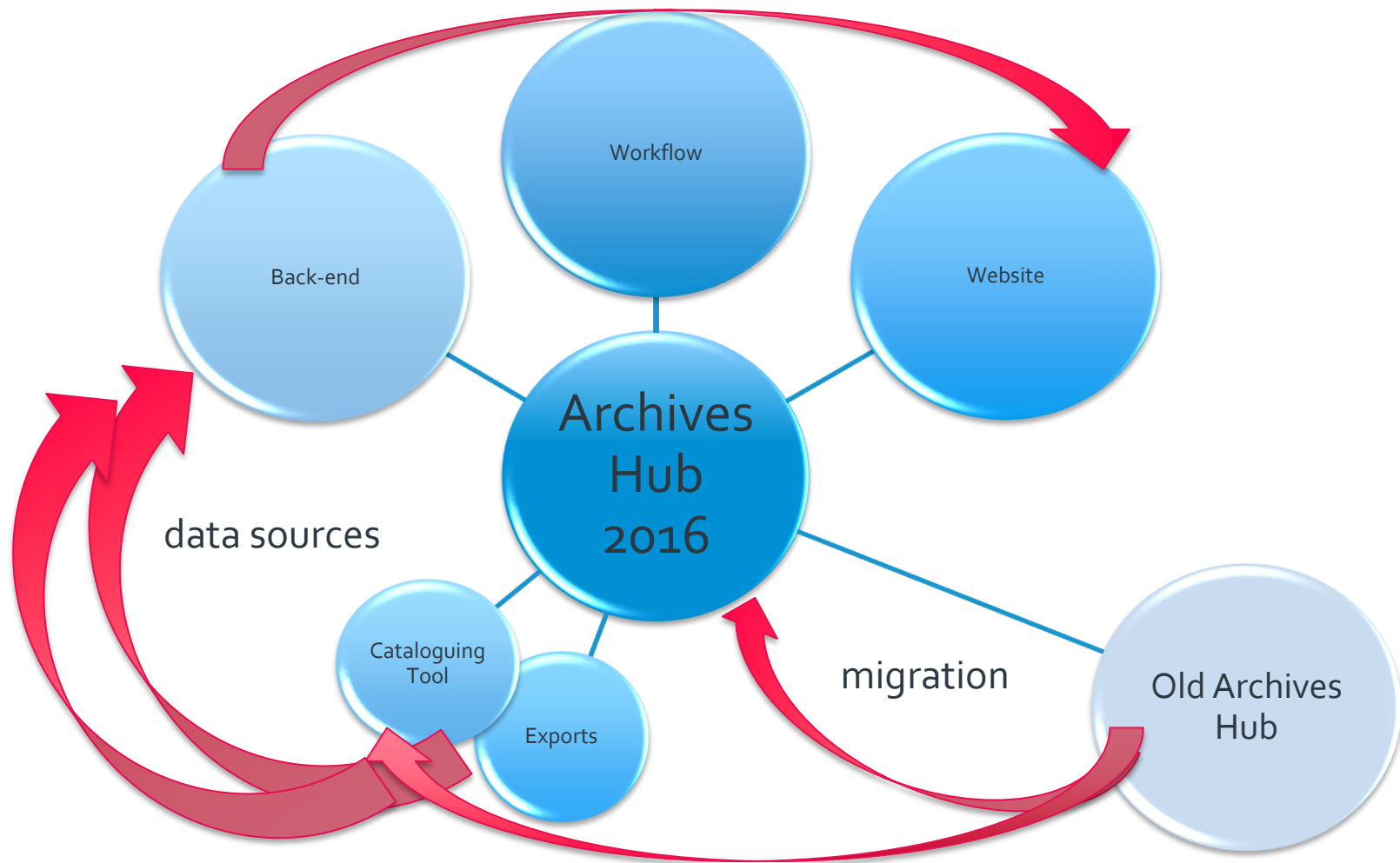
**320 contributors**

**Any kind of  
archive**

**Contribute as and  
when they wish**

**Collection level  
and multi-level**

**Use our  
cataloguing tool  
or export from  
systems**



- » Migrate descriptions from old system --- normalise
- » Workflow for new system --- repeatable processing
- » Rigorous treatment of data --- consistency, structure
- » Sustainable model
  - › Entities: Archives, Repositories, Names
- » Take data in automated way & at scale
- » Deal effectively with complex hierarchy & revisions

» Check for:

- › Valid EAD
- › Hub requirements
- › Content

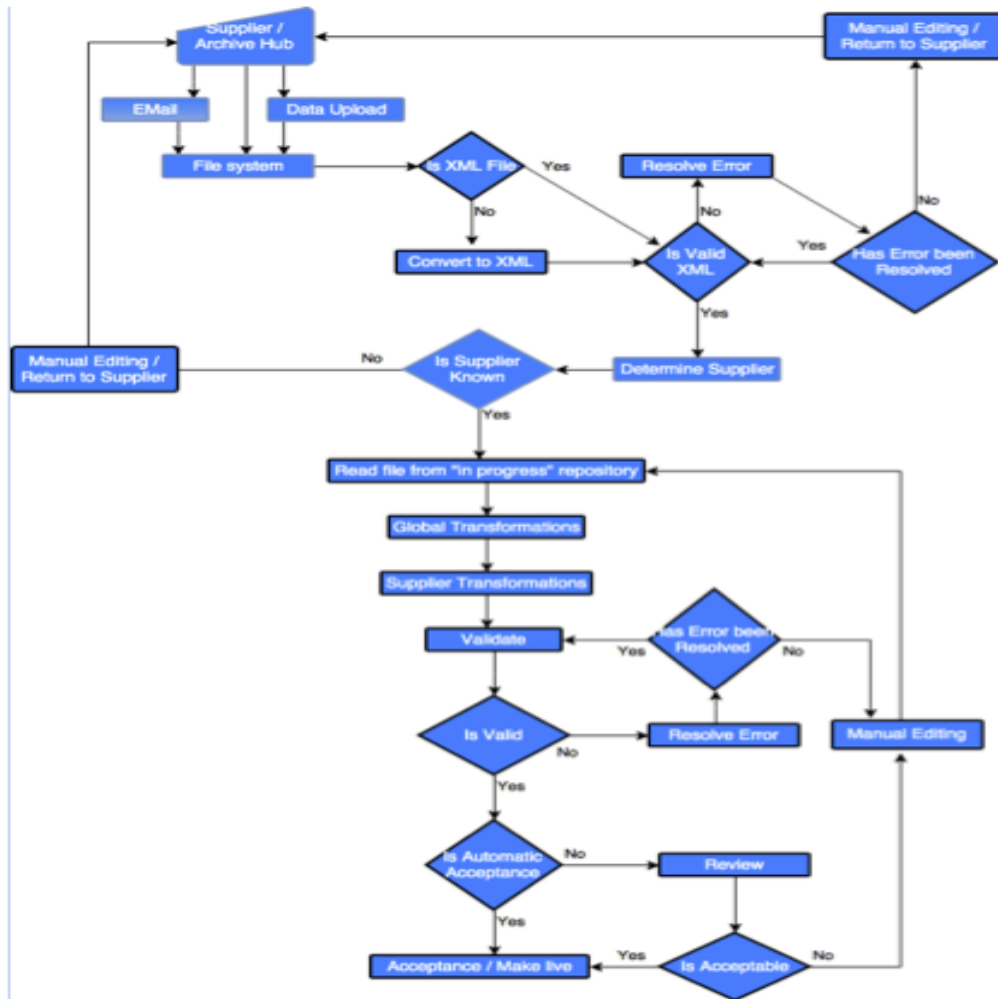
» Revisions --- version control issues

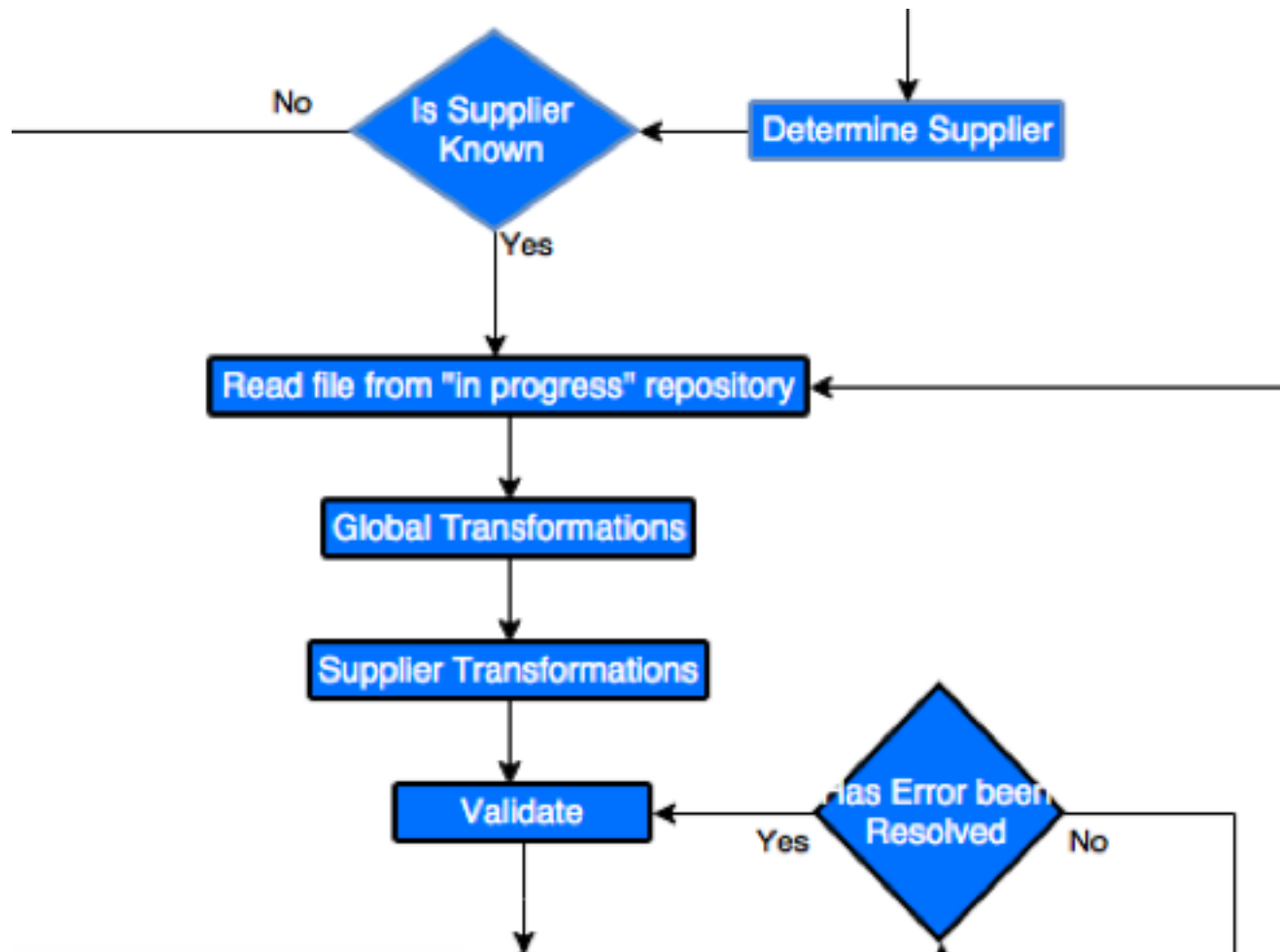
- › More than one version of a description
- › Which is the master copy?
- › Can change title, reference, filename

**Solution needs to  
work for wide variety  
of descriptions**

## Archives Hub Workflow

Knowledge  
Integration:  
CIIM

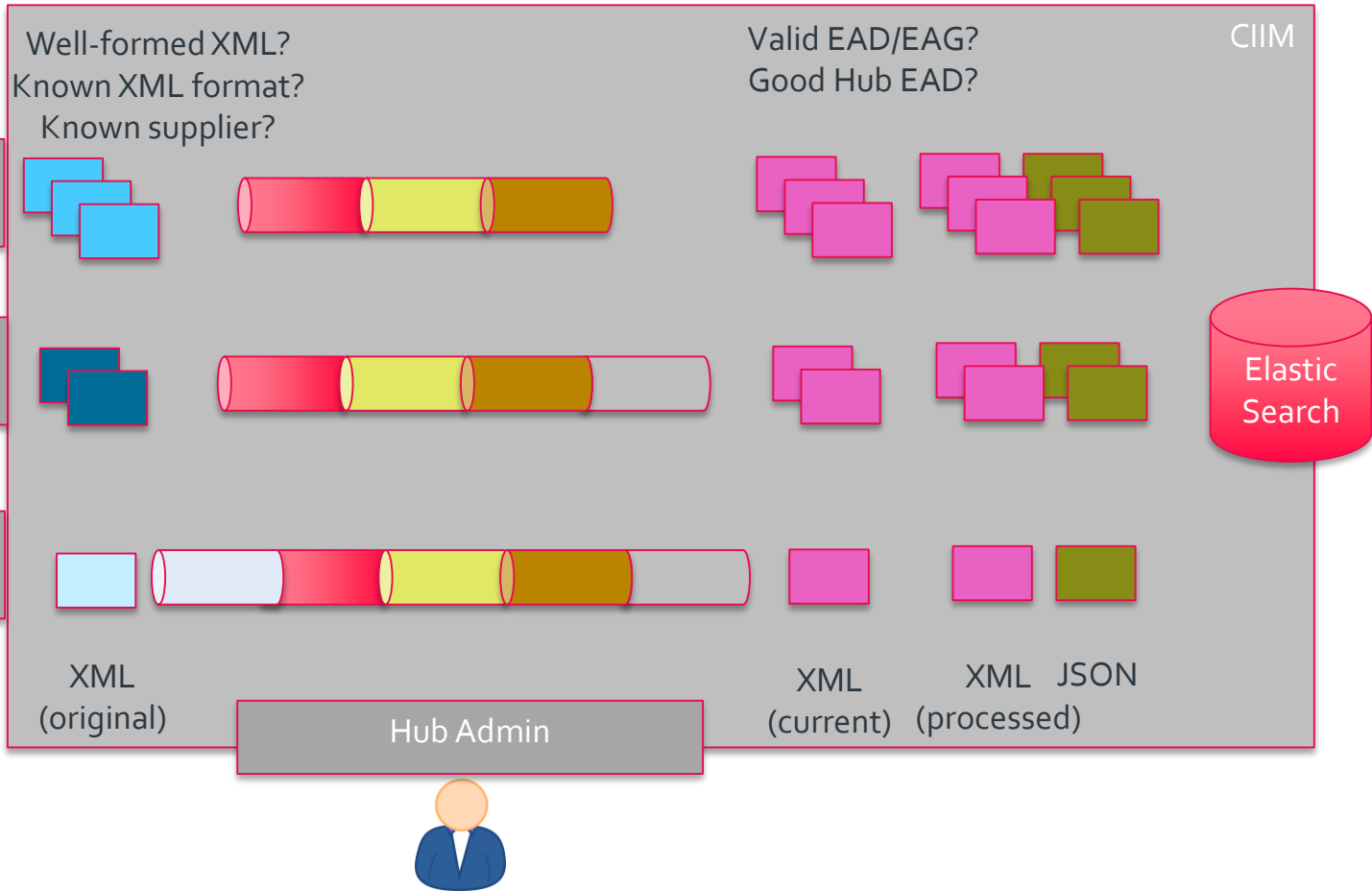


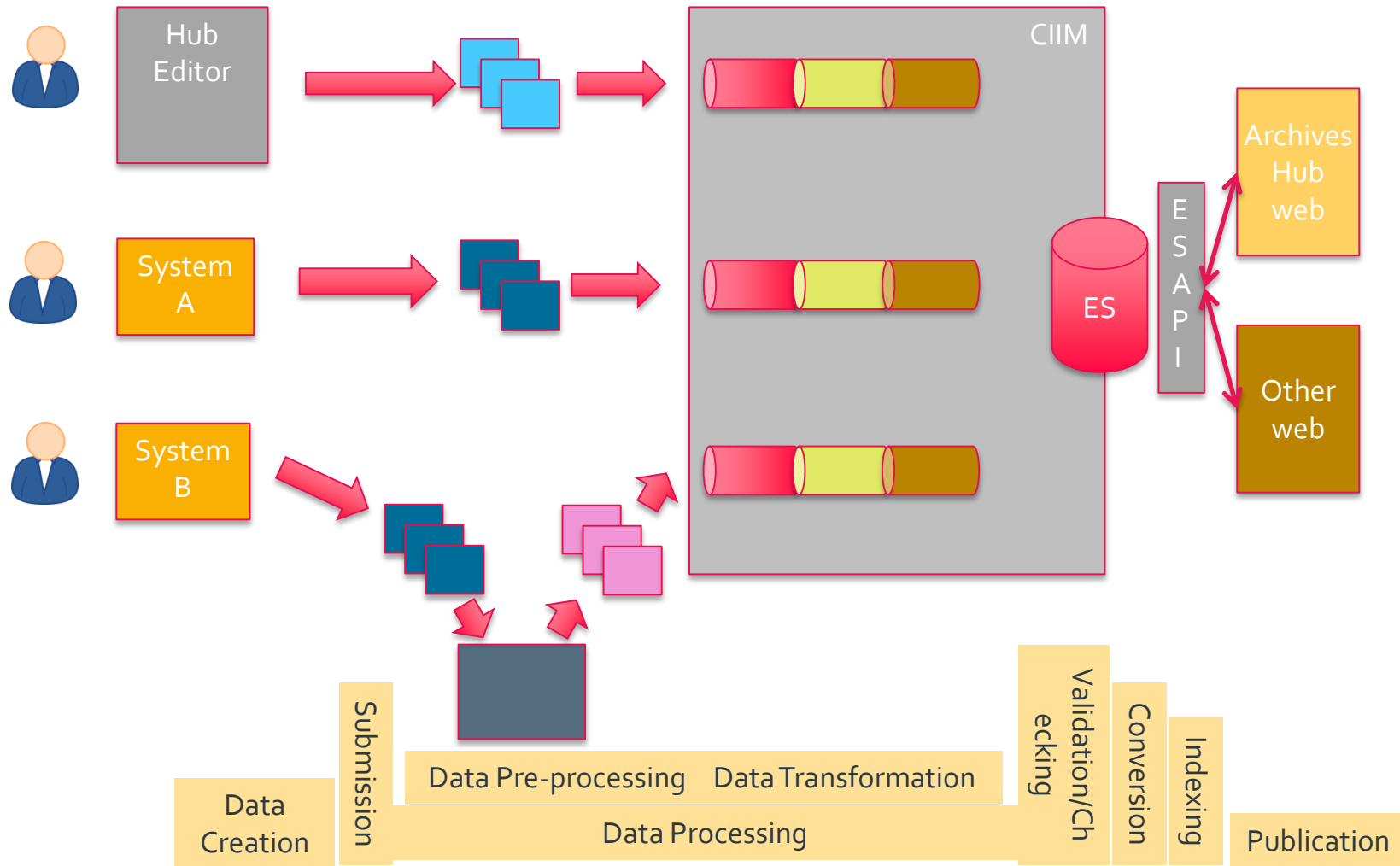


- » What are we looking to achieve?
- » What are the required data conditions?
- » What is the rationale for them?
- » How can they be implemented consistently?



- » References must be unique
- » Repository name must not include the archon code
- » Level values from a controlled list
- » Language must be provided at the top level
- » Access conditions must be provided at the top level
- » Normalised dates must use the correct ISO format
- » URLs must use http://
- » [Title below 500 characters]







Browse



Data Manager



Organisations



Users



Tasks



Endpoints

[Files](#) [System Files](#) [Transformation Sequence](#) [Validations](#) [Settings](#)

## Upload Data File

Organisation:

Select File:

[Browse...](#)[Upload](#)

## Refine Results

Order by:

[Date](#)[Filename](#)[Title](#)[! desc ▾](#)

The number of files to display

Workflow Status:

- ✓ All
- Errors
- Conflicts
- Created
- Manual Review
- Published
- Archives Hub to Resolve

Showing 1 - 10 of 116985 Files

⏮ ⏪ 1 2 3 4 5 ... 11699 ⏩ ⏭

[Select All](#)[Archives Hub to Resolve](#)[Contributor to Resolve](#)[Publish](#)[Repeat All Transforms](#)[Repeat Global Transforms](#)[Reprocess Original](#)[Send to ES](#)[Validate](#)

0 Files Selected

<input type="checkbox"/>	<b>EAD</b>	<b>Llansamlet Ecclesiastical Parish Records</b> P 60 CW westglam P 60 CW.xml	Published 14-May-2018 18:34   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>	<b>EAD</b>	<b>Loughor Ecclesiastical Parish Records</b> P 112 CW westglam P 112 CW.xml	Published 14-May-2018 18:34   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>	<b>EAD</b>	<b>Swansea, St Mary Ecclesiastical Parish Records</b> P123CW westglam P 123 CW.xml	Published 14-May-2018 18:13   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>	<b>EAD</b>	<b>Pennard Ecclesiastical Parish</b> P 117 CW westglam P 117 CW.xml	Published 14-May-2018 18:13   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>	<b>EAD</b>	<b>Llwynderw Ecclesiastical Parish Records</b> P 310 CW westglam P 310 CW.xml	Published 14-May-2018 18:13   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>	<b>EAD</b>	<b>Morrison Ecclesiastical Parish Records</b> P 312 CW westglam P 312 CW.xml	Published 14-May-2018 18:13   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>	<b>EAD</b>	<b>Manselton Ecclesiastical Parish Records</b> P 311 CW westglam P 311 CW.xml	Published 14-May-2018 18:13   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>



Browse



Data Manager



Organisations



Users



Tasks



Endpoints

Files

System Files

Transformation Sequence

Validations

Settings

## ⬆ Upload Data File

Organisation:

Select File:

## 🔍 Refine Results

Order by:

The number of files to display

Workflow Status:

## Showing 1 - 10 of 34 Files

0 Files Selected

<input type="checkbox"/>		<b>EAD</b>	<b>Wynn Hall Mss</b> WYNN bangor WYNN.xml	<b>6 Errors</b> Error Validation 08-Jan-2018 09:04   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>		<b>EAD</b>	<b>Huddersfield Branch of the British Federation of Women Graduates</b> WGH huddersfield WGH.xml	<b>4 Errors</b> Error Validation 18-Jan-2018 15:57   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>		<b>EAD</b>	<b>Ted Hughes Network Archive</b> THN huddersfield THN.xml	<b>1 Errors</b> Error Validation 18-Jan-2018 15:57   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>		<b>EAD</b>	<b>huddersfield</b> tfl.xml	<b>1 Errors</b> Error Removing 14-May-2018 15:37   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>		<b>EAD</b>	<b>St. Andrew's Society Archive</b> STA huddersfield STA.xml	<b>1 Errors</b> Error Validation 18-Jan-2018 15:57   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>
<input type="checkbox"/>		<b>EAD</b>	<b>Gnoll Estate Collection</b> RISW Gn westglam RISW Gn.xml	<b>1 Errors</b> Error Validation 11-Sep-2017 12:18   <a href="#">View Audits</a>	<a href="#">Actions ▾</a>

[Select All](#)[Archives Hub to Resolve](#)[Contributor to Resolve](#)[Publish](#)[Repeat All Transforms](#)[Repeat Global Transforms](#)[Reprocess Original](#)[Send to ES](#)[Validate](#)

0 Files Selected

**EAD**

Wynn Hall Mss  
WYNN  
bangor WYNN.xml

**6 Errors**

Error Validation

08-Jan-2018 09:04 | [View Audits](#)[Actions](#) ▾

## Status: Error Validation

**6 Errors**

**XML Validation Error, Location: /ead[1]/archdesc[1]/dsc[1]/c[3]/c[2]/did[1]/unitid[1]**

For the persistent or current unitid, element unitid attribute identifier value MUST be unique. [Reference(s): WYNN/736] [identifier: WYNN/736]

**XML Validation Error, Location: /ead[1]/archdesc[1]/dsc[1]/c[3]/c[2]/c[1]/did[1]/unitid[1]**

For the persistent or current unitid, element unitid attribute identifier value MUST be unique. [Reference(s): WYNN/736] [identifier: WYNN/736]

**XML Validation Error, Location: /ead[1]/archdesc[1]/dsc[1]/c[3]/c[5]/did[1]/unitid[1]**

For the persistent or current unitid, element unitid attribute identifier value MUST be unique. [Reference(s): WYNN/808] [identifier: WYNN/808]



Files

System Files

Transformation Sequence

Validations

Settings

## Groups

AtoM

2

Basic

2

BasicAdd

3

BasicWithAdd

4

BritishLibrary

4

Calm

5

CalmWales

6

Manchester

4

Pete

Test

## Organisations

Global Organisation for Workflow (Global)

16

Aberdeenshire Museums Service

Aberystwyth University Archives

Ahmed Iqbal Ullah Race Relations Resource Centre

AHRC Research Centre for the History and Analysis  
of Recorded MusicAlexander Library of Ornithology, University of  
Oxford

Amaravati Buddhist Monastery Archives

Anthony Burgess Foundation Archives

## Global Transformations

1 EAD

TidyThemed

tidy-themed.xml



Global Organisation for Workflow

Organisation

18-Jul-2017 15:23

(File) Last Updated



Disable



2 EAD

FixNormalisedDates

fix-normalised-dates.xml



Global Organisation for Workflow

Organisation

30-Nov-2016 18:04

(File) Last Updated



Disable



3 EAD

RepairExtref

repair-extref.xml



Global Organisation for Workflow

Organisation

07-Dec-2017 18:11

(File) Last Updated



Disable



4 EAD

FixLangcodes

fix-langcodes.xml



Global Organisation for Workflow

Organisation

28-Nov-2016 15:26

(File) Last Updated



Disable



5 EAD

DedupAll

dedup-all.xml



Global Organisation for Workflow

Organisation

29-Jul-2016 10:40

(File) Last Updated



Disable



6 EAD

TidySourceAndRules

tidy-ca-source-rules.xml



Global Organisation for Workflow

Organisation

17-Apr-2018 15:58

(File) Last Updated



Disable



7 EAD

TidyEmphAlt

tidy-ca-emph-attrs.xml



Global Organisation for Workflow

Organisation

30-Nov-2016 18:21

(File) Last Updated



Disable



8 EAD

TidyGeogName

tidy-ca-geogname.xml



Global Organisation for Workflow

Organisation

26-Apr-2017 17:00

(File) Last Updated



Disable



9 EAD

NormaliseLevels

normalise-levels.xml



Global Organisation for Workflow

Organisation

16-May-2016 15:11

(File) Last Updated



Disable



10 EAD

RepairRepoName

repair-reponame.xml



Global Organisation for Workflow

Organisation

11-Dec-2017 12:59

(File) Last Updated



Disable





**Contributor**  
“Local”, specific  
to Contributor

**Group**  
Common to  
multiple  
Contributors

**Global**  
Common to  
all  
Contributors



» Same model applies to all data formats:

› EAD (archive descriptions)



36,510  
ARCHIVE

› EAG (repository descriptions)



335  
REPOSITORY

› EAC (people and organisations)



141,291  
AGENT

- » Descriptions often created to work within one system
- » Designed for one interface
- » Constraints of what is achievable within one system
- » Data may not be suitable for an aggregator, for interoperability, for different systems and environments
- » ...but it is hard to modify systems or change cataloguing practices
- » Need a compromise

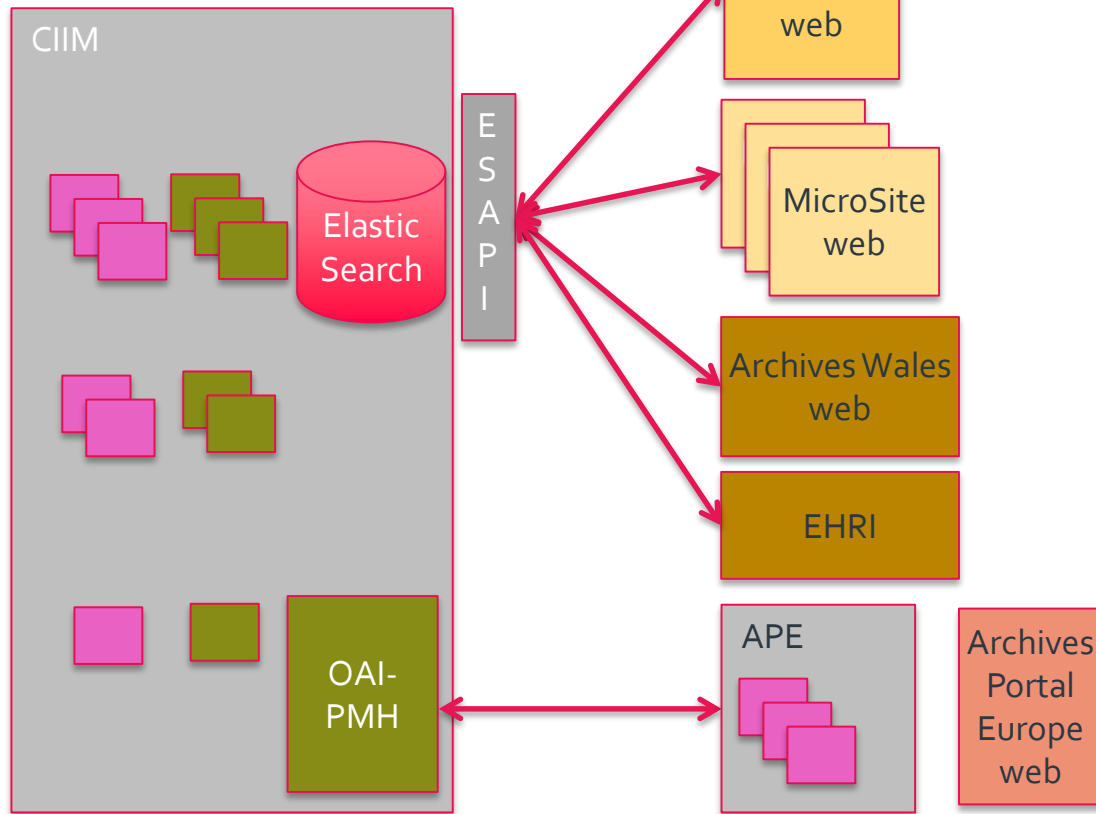
- » Missing references
- » Duplicate unique references
- » Missing mandatory fields
- » Titles that are just a name...or they are a page long
- » Cutting and pasting resulting in Jean HÃfÆ'Ã†â€™
- » Dates in all manner of formats: [ ( c ca c. ? - / 1890-1831
- » Data that isn't structured




- » Customised fields
- » Printouts first
- » Levels values in welsh
- » Unesco subject terms that aren't from the unesco thesaurus
- » Place as name, name as subject, subject as genre, subject as place...
- » Ghost levels – physical versus intellectual



- » An archivist typically thinks about describing an archive in a way that a human user can understand, accessed from their reading room or website; we think about it as structured data for humans and machines, within a global context




Archives Hub

All
magic
Search

Search help

Advanced Search

Search Results

Filter
Applied Filters
Level
collection
Remove
Filter by
Holding Repository
Level
collection (65)
section (248)
item (116)
Creator
Subject
Date
Digital Content

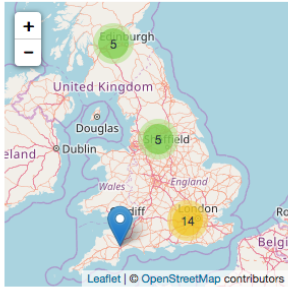
Relevance
Show 20
Expanded mode

64 Archive Collections
1 Themed Collections
0 Repositories

< Previous
1 2 3 ... 4
Next >

1 Notes on Magic
19th Century
Show on map
2 Treatise on Magic
1854-1857
Show on map
3 Batak Pustaha [Magic Book]
n.d.
Show on map
4 'Religion, myth and magic in Tanqala'

Show Map



Filter

Filter by

Holding Repository ▾

Level ^

collection (36974)  
section (408508)  
item (934643)

Creator ▾

Subject ▾

Date ^

0001 - 0999 (353)  
1000 - 1099 (131)  
1100 - 1199 (396)  
1200 - 1299 (1289)  
1300 - 1399 (2314)  
1400 - 1499 (2587)  
1500 - 1599 (8458)  
1600 - 1699 (30340)

Relevance ▾

Show 20 ▾

1382790  
Archive Collections

64  
Themed Collections

< Previous

1

2

3

...

69140

Next >

1  
Archive Collections

Society for Promoting the Training of Women

1859-1991

Show on map

2  
Archive Collections

Personal Papers of Maud Chapman

1898-1905

Show on map

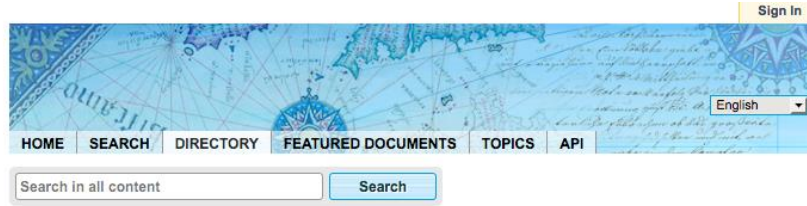


## » OAI-PMH

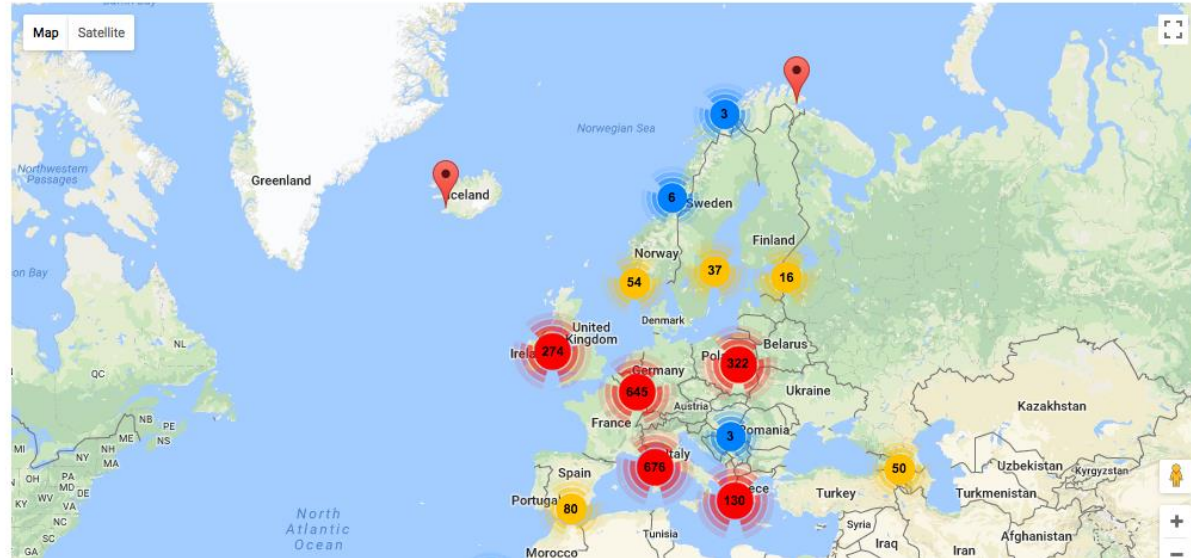
- › Archives Portal Europe
- › Cendari

## » Elastic Search API

- › Archives Wales
- › Micro sites
- › European Holocaust Research Institute



## Directory



Enter the URL of your repository:

OK

Enter the URL of your repository:

Select a set:

Select a metadata prefix:

Select harvesting method:

Select an interval for your harvesting:


Harvest only on weekends?

Set last harvesting date (if existing) -  
dd/MM/YYYY:

Select a user profile:

Activation status:

OK



Search Results

Archives Hub: University of Salford

Help

All

▼

Enter search...

Search

▼

Advanced Search

Filter

Filter by

Level ▼

Creator ▼

Subject ▼

Date ▼

Digital Content ▼

Date (Descending) ▼

Show 20 ▼

Expanded mode ☐

9749

Archive Collections

< Previous

1

2

3

...

488

Next >

1

Roy Newsome Brass Band Archive

Sun Life and Stanshawe Band

2010

2

Roy Newsome Brass Band Archive

Wingates Band

2010

3


Roy Newsome Brass Band Archive

Clayton Aniline Works Band

2010



**Jisc Archives Hub** Search Help

 Archives Hub Place ▼ wales Search

[Search help](#) Advanced Search

Search Results

**Filter** <<

**Applied Filters**

**Holding Repository**

- Swansea University Archives Remove
- Bangor University Archives Remove
- West Glamorgan Archive Service Remove
- Cardiff University Archives Remove
- Aberystwyth University Archives Remove




**Filter by**

**Holding Repository** ^ (Showing top 20 of 49)


- ✓ Swansea University Archives (457)
- ✓ Bangor University Archives (335)
- Sedgwick Museum of Earth Sciences, University of Cambridge (23)
- St John's College Library, University of Oxford (22)
- ✓ West Glamorgan Archive Service (16)
- London School of Economics Archives (15)
- Modern Records Centre, University of Warwick (14)
- Cambridge University Library (10)
- National Co-operative Archive (10)
- Bodleian Library, University of Oxford (9)
- University of Brighton Design Archives (7)
- School of Oriental and African Studies ...


Relevance ▼ Show 20 ▼ Expanded mode ☐


**818** ? **0** ? **0** ?

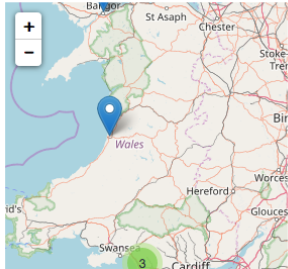
 Archive Collections  Themed Collections  Repositories

< Previous **1** **2** **3** ... **41** Next > Show Map ☐

**1**  **Enoch Robert Gibbon Salisbury (1809-1890) Miscellaneous Manuscripts and Papers**  
c.1840 - [ongoing]  
[Show on map](#)

**2**  **Salisbury Library: Additional Manuscripts**  
1886 - [ongoing]  
[Show on map](#)

**3**  **Marjorie Chaplin, Left Book Club**



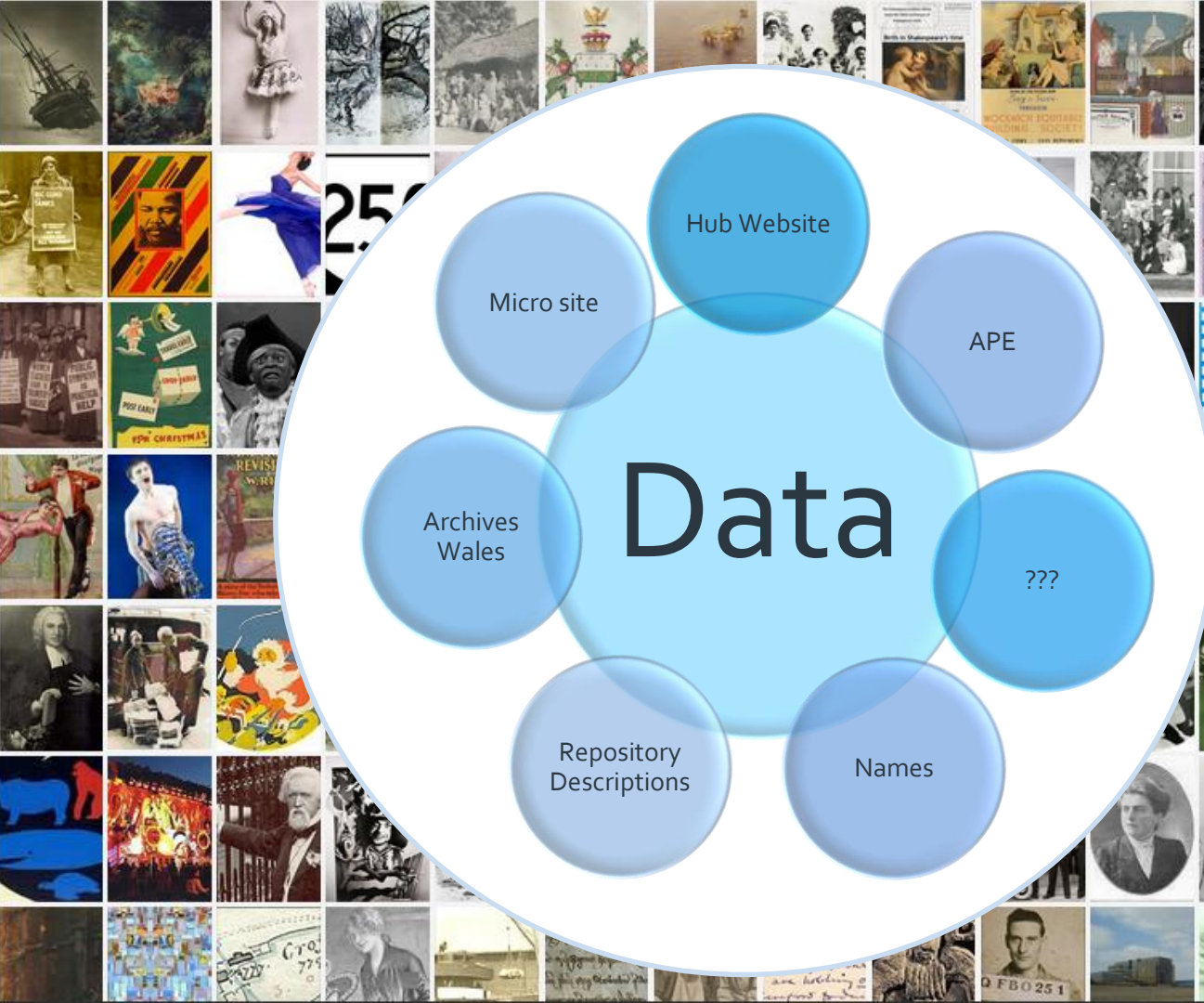
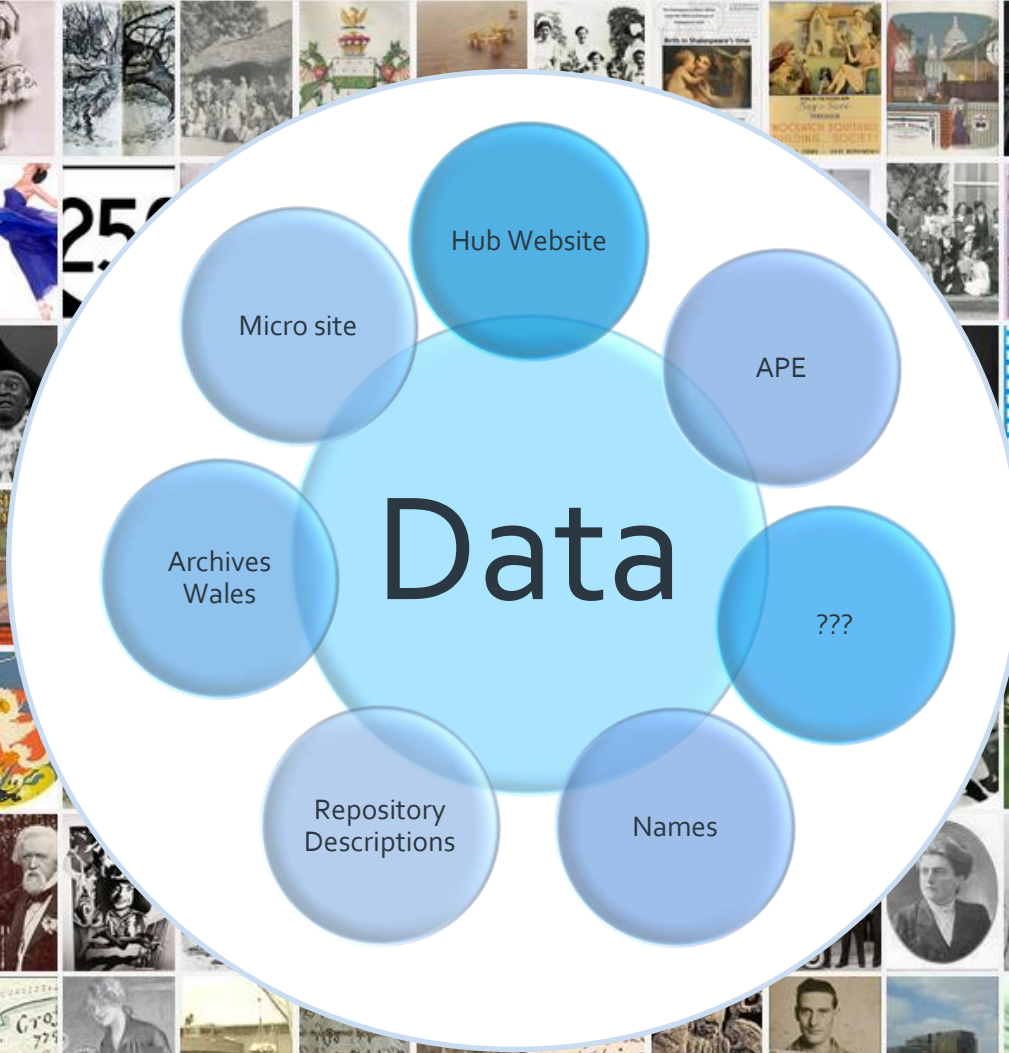












Jane Stevenson  
Archives Hub Service Manager

jane.stevenson@jisc.ac.uk

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

archiveshub@jisc.ac.uk



Except where otherwise noted, this work is licensed under CC-BY-NC-ND