Zoom chat between Kieran and Brenda today 11:30
Topic: Processes

Step 1:

Divide papers into categories that have a correlation. This allows for batch cataloguing. This was done manually – staff members were given A4 pages with boxes. They had to assign each document a unique number. (When you see one of the templates you will see where this is coming from).
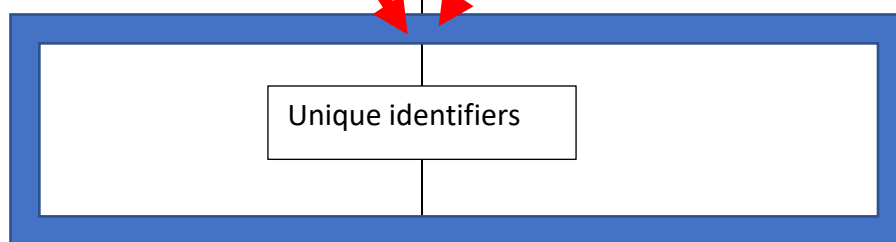
Step 2:
Staff member enters the data into an XL spreadsheet.

Step 3:
Quality checks
Check that the transcribed record that is on the spreadsheet is matching with the artefact.

Unique identifiers

Step 4:
Take details in the spreadsheet. The software used is MarcEdit. Align your tabular spreadsheet to a particular column:
e.g. Author -> Marcfield 100

Data would be manipulated here. Standardisation on things like:

Surname,[space here] Firstname

Done in parallel

Artefacts are digitised.

File name map to unique identifier that was in the initial record.
Consistent structures as regards file types: TIFFS used.

Would look at using authority records here e.g. Library of Congress (LOC).
There could be variation of name but it links back to the one standard authority.
Library of Congress subject headings as well [so in that regard they are not that far away from a standard]

The MARC file contains the URL so have to make sure that areas on the web are available and that the data is there.

Step 5:
The MARC file is entered into the KOHA library management system. (KOHA is not ideal as the modification of records on system is hard to do).

Step 6:
From KOHA a MARC xml file is exported which can be ingested into the DRI.

This is important – no changes can be made after upload to DRI they have to be done in advance of uploading. Making changes after upload is bad research practise.

Graves papers structure of title

Letter from	Letter To	Date