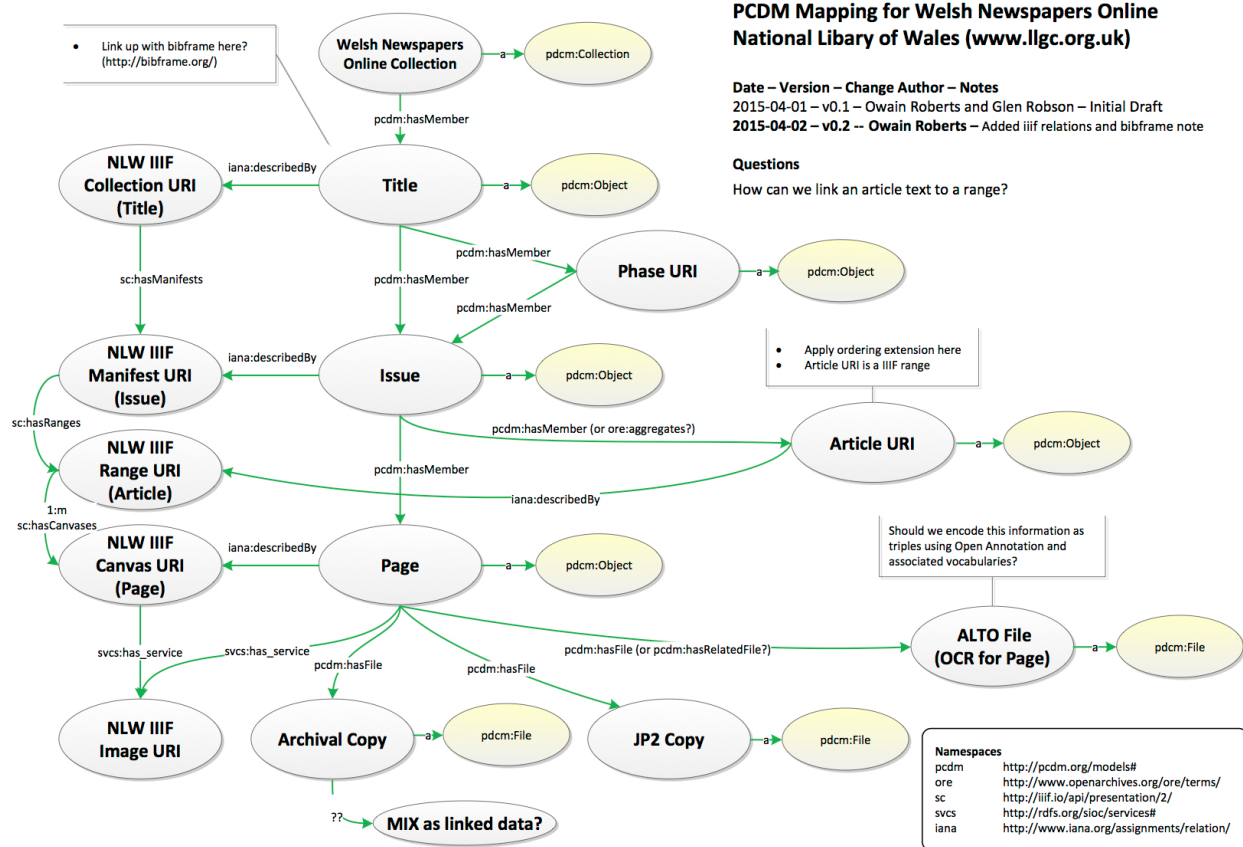


PCDM Mapping for Welsh Newspapers (NLW)

This page hopes to document the first attempt at mapping the National Library of Wales Newspaper content to PCDM and IIIF. A diagram of the Newspaper in PCDM is below:



Note:

- Title: is a Newspaper Title which has a record in our MARC catalogue.
- Phase: is a physical unit of handling (for digitisation) and is usually a set of issues physically bound together. We use this information to manage batches for digitisation and is not displayed to users.
- Issue: an issue of a Newspaper, this has a ISO issue date as metadata
- Article: a newspaper article, this can span Pages and can have multiple columns on one page. There are many articles on a single page.
- Page: a physical page of a Newspaper and also a container for Scanned image of a page.
- Archival Copy: Archival TIFF held in a HSM near line storage. Referenced over HTTP from Fedora
- JP2: reference version of the page currently stored as a managed datastream in Fedora
- ALTO: OCR Text, Article Boundaries and Coordinate information generated from the TIFF.

- We've added IIIF classes and relations into the above diagram to specify the OCR text annotations on a page and the related article metadata.

Questions:

- In the [Portland Common Data Model](#) it specifies the rdfs:label on the File can be repeatable. We could only thing of a file having a single Filename, is there an example where to labels for a File might be required.
- We haven't put it in the diagram but if a Manuscript had two orders one physical order (the order the physical material is in pre-scanning) and logical order (maybe the font covers have been moved from the back to the front). Would you have 1 object for the Manuscript and two member objects for each order? The proxies could then link the the Order Objects.
- This is probably a IIIF question but we struggled to link the text of an article with the article object. We modeled a Newspaper Article to a range (as it can cross pages) but we couldn't see how we could add an annotation to a range as an annotation seemed to be limited to a Canvas.

Comments welcome!