CONWAY HALL

Platform Scoping

Introduction

Digitisation covers many different approaches to making physical objects and collections available online. A simple form of digitisation has already taken place at Conway Hall with scans of the artwork and pdf files of the Ethical Record made available to the public using our current website platform, Wordpress. However digitisation within the GLAM[1] sector emphasises not only providing access to high quality images, established as best practice, but also the provision of high quality metadata. There are a number of different platforms cultural institutions have used in order to host their digital collections; these are usually dictated by their own institutional requirements/limitations:

- The types of collection
- The contents of the collection
- Core audience/s
- Cost
- Internal technology infrastructure, actual or potential
- Internal digitisation know-how

Our overall institutional requirements/limitations are as follows:

- Requirements
  - The collections are varied but are predominantly document based and are a mix of archival and library material
  - The content of the library collections is somewhat wide-ranging but has particular focus on philosophy, history, freethought, humanism and religion. Our archives are dedicated to the processes and activities of the Conway Hall Ethical Society, The Sunday Concert Programme and those of the National Secular Society
  - Our digitisation pilot project (Architecture and Place) prospective audiences are defined initially as members of the Conway Hall Ethical Society the general public, locals, and architects and architecture students.
    - An audience we are more broadly interested in reaching with future projects are researchers. Please note that neither of these lists are exhaustive.
- Limitations
  - As a charity we are required to be frugal and are therefore unlikely to invest in mass digitisation projects or expensive licensing agreements
    - Unlike academic libraries and archives (which do a large amount of digitisation) we do not have the human resources or capital funds to support 'complete' digital humanities level digitisation comprising of high quality

---

[1] Galleries, Libraries, Archives and Museums

images, html produced from fully transcribed texts encoded into TEI XML with high level entity authority control
- There is funding for initial development but it is not large and there will be limited funds to maintain the platform going forward
- There is one dedicated staff member with knowledge of digitisation but the majority of IT support is external

Prior to deciding on which platform is most appropriate for your collections and audiences, a decision can be made around whether to focus on assessing Open Source(OS) solutions &/or consider proprietary software. Each option has its own pros and cons (see fig.1 below).

	Pros	Cons
Proprietary	Support, training, setup easier, up-front notice of changes	Ongoing and up-front costs, locked in to vendor, not extensible, slower updates or bug fixes, over-selling of features
Open Source	Extensible, cheaper, open licensing, interoperable, community support (people in your situation giving advice rather than call centre or sales staff)	Support, training, setup more difficult

Figure 1

Open Source software allows you to try something to see if it fits without having to spend too much money and it can help clarify some problems you might not know exist. OS has a community of people in similar situations so you can ask for advice that will be relevant rather than receive a pre-scripted response. OS is more likely to be interoperable and have some way of getting your data out into another system. OS usually has a lot more plugins to use compared to proprietary software and there are usually more frequent releases compared to proprietary.

For Conway Hall we focussed on Open Source software alone predominantly because of the ongoing licensing costs of proprietary software although other factors, such as open formats, interoperability and active developer and user communities were also considered important.

Another decision that can be made before considering platforms is to assess whether an independent platform will be of more benefit than a platform that integrates with our current catalogue. Our cataloguing system, Heritage, does not offer the functionality to allow us to consider it as an option for our digitisation platform.

Ultimately we want a platform that: is easy to use and re-configure; its features conform to best practice; is stable; allows for us to create exhibitions with different items within the larger collection/s and can scale up. As our collections are made up of a wide variety of media we need a platform that is flexible and suited to supporting a number of formats.

Functionality Requirements

The following functionality requirements were developed in a number of discussions between Library and other Conway Hall staff and from research undertaken for this project. They include required functionality as well as optional aspects.

Required

- Ease of use
- Handling of Entity Authority/controlled vocabularies
- Metadata support, ingest and export facilities
- Support (user and developer communities)
- Support of learning resources and any online learning activities
- Handling of external links
- Cost
- Search API(so that the digital collections are searchable from the Conway Hall website)
- Full text searching (within PDFs especially)
- OAI-PMH compliance[2]
- Persistent identifiers (URIs)[3] for digital objects
- Support large format images , video, audio, pdf, word documents, ePub and so on
- Create exhibitions

Optional

- Citation export/social media shareability
- Crowdsourcing functionality (e.g. transcription)
- User contributed material, tags etc

Initial Platform Comparisons

There are a number of open source platforms that are used in the GLAM sector to digitise collections. However, several of them we knew from the outset were unlikely to be good matches for our requirements, due to difficulty of use or not being a good match for our collections, and thus were not evaluated thoroughly against the functionality requirements outlined above.[4] These included:

- Institutional repositories such as:
  - E Prints
  - D Space
  - Fedora

---

[2] Open Archives Initiative Protocol for Metadata Harvest is used to expose your metadata to aggregators and other harvesters.

[3] URIs (Uniform Resource Identifiers) that are persistent minimise link rot (https://en.wikipedia.org/wiki/Link_rot).

[4] These platform candidates were identified in the literature (see references at the end of document), a variety of forums, and in email discussions with Conal Tuohy, a former colleague and independent consultant working in the Digital Humanities sector with libraries and archives in Australia.

- o Invenio

Institutional repositories are primarily intended to store and disseminate the research outputs for academic and research institutions. They are predominantly text heavy and require significant IT development and ongoing support. New York Public Libraries uses Fedora for part of their digital collections but has spent substantial funds in supporting their creation and developing tools to use them.

- Greenstone

Greenstone is a digital library system developed in New Zealand at the University of Waikato in cooperation with UNESCO. It has had a relatively wide uptake within communities connected to non-profit organisations and NGOs who work with UNESCO. It is not particularly user friendly and does not lend itself well to building exhibitions.

- Kete

A community developed digital assets management system that currently has little support from its previous developer/user groups and is not widely used outside of New Zealand.

- Drupal

Drupal is an open source content management system that has a wide user base but does not lend itself well to easy administration and is not recommended by a colleague who has used it.

- XTF (eXtensible Text Framework)

Developed for the creation of text heavy collections and difficult to administer XTF is therefore an unsuitable digital platform candidate for our mixed format collections.

Please note that this is not an exhaustive list of open source content management systems; or of digital asset management, digital library or digital archive platforms. Those considered were the most well-known candidates or platforms we or colleagues had personal experience with using.

Another contender that is not open source that was included for preliminary consideration was Google Cultural Institute. It was included as the recommendation of one of the Society's trustees and the CEO. Google Cultural Institute is used by many large, well known cultural heritage institutions, such as the British Museum, to digitally expose their collections and exhibitions. However, neither these institutions nor Google seem to promote these digital collections widely. The focus, it appears, is on providing content to Google's products, specifically Google Maps and Google Now.

As Google Cultural Institute is not core to Google business there is a risk that development of the product will be dropped, for example, such as that for Google Glass. In addition, Google Cultural Institute provides limited analytics and there is some question around the copyright of metadata. Given these limitations we will not be considering this Google product as the sole platform for our digital collections and exhibitions but will consider it when creating a virtual tour through Conway Hall and for specific exhibitions.

Platform Evaluation

The following platforms are those that met an initial cursory appraisal and were subject to fuller evaluation against the functionality requirements outlined above.

- Omeka

Omeka is a content management system developed specifically for online digital collections and was especially recommended as a suitable candidate for our digitisation projects.[5]

- AtoM (Access to Memory)

AtoM was primarily developed as an archival description system but has been used as a digitisation platform and mentioned as a possibility on a UK archives email list.

- Islandora

Islandora is a software framework for the management and discovery of digital assets and makes use of Fedora, Drupal and other software applications.

Each candidate was assessed against the functionality requirements (see Appendix A for full evaluation notes); a summary of this process can be seen below. Evaluation was undertaken through some practical experience (Omeka only) but primarily through software documentation, evaluations by others, comparison articles, and experience with collections built on the candidate platforms. One disclaimer is that weightings given for each criterion are subjective and entirely bound by the limits of the research undertaken.

---

[5] "…Omeka is definitely the best product of that type that I know of. It has a decent sized user base and a developer community with funding. So you can't go too far wrong selecting it I think. It also has some input/output mechanisms for migration if it ever comes to that," Private email from Conal Tuohy, 19 January 2016.

Platform Evaluation Summary

Platform	Description	Pros	Cons	Score	Notes
Omeka **https://omeka.org/**	CMS/web publishing platform Designed for digital collection dissemination	Ease of use Exhibition functionality	Metadata handling Not able to support large scale digitisation of texts to a digital humanities level	**40**	Includes Neatline[6] a suite of plugins for Omeka to build geotemporal exhibits/narratives from your digital collections "items-in-a-narrative"[7]
AtoM **https://www.access tomemory.org/en/**	Archival description platform	Description standards-centric Widely used within archives Edited via webpages	Does not support large image formats No exhibition functionality Very easy to delete an entire collection Setting up collection from scratch is time consuming	**29**	For archival description and access but can be used by libraries, museums etc Multi-repository 'union list' accepting descriptions from any contributing institutions – could partner with Bishopsgate?
Islandora **http://islandora.ca/**	Digital repository system	OCR functionality Browser compatibility tested Version control Can support mass digitisation	Complicated to install, configure and likely use Likelihood we'd require a hosted solution with associated costs Sustainability risk due to dependence on other OS software No exhibition functionality	**36**	Models and captures complex relationships between digital objects Linux based Documentation is developer driven

---

[6] http://neatline.org/about/
[7] http://omeka.org/forums/topic/omeka-vs-drupal-why-might-one-use-omeka-instead#post-4950

Conclusion

Evaluation was a difficult and lengthy process but Omeka came out ahead primarily because it is the best fit for our needs within the limitations (of funding and staffing) we operate under.

The three candidates that were fully evaluated were, at various times, considered the strongest contender: AtoM for its standards-focus and the clarity of the websites built with it; Islandora because of its range of digitisation possibilities and conceptual model. Omeka however is very user-focussed (cultural heritage staff and end-user) and is therefore straightforward and easy to use. It was developed with the creation of online exhibitions in mind which was one of our primary considerations and it better matched our required criteria than any other candidate (33 vs. 27 [AtoM] vs. 27 [Islandora]).

Implementing Omeka will begin following signoff from the CEO.

References

Das, Amiya. "Comparing Open Source Digital Library Software: Special Reference to DSpace, EPrint and Greenstone." *International Journal of Advanced Research in Computer Science and Software Engineering* 5(7): 70-3. Accessed November 25, 2015. http://www.ijarcsse.com/docs/papers/Volume_5/7_July2015/V5I7-0132.pdf.

Kökörčený, Michal, and Bodnárová, Agáta. "Comparison of Digital Libraries Systems." Paper presented at the 9th WSEAS International Conference on Data Networks, Communications, Computers, Faro, Portugal, November 3-5, 2010.

http://www.wseas.us/e-library/conferences/2010/Faro/DNCOCO/DNCOCO-16.pdf

Pirounakis, George, and Nikolaidou, Mara. "Comparing Open Source Digital Library Software." Accessed November 25, 2015. http://www.thailibrary.in.th/wp-content/uploads/2011/05/Comparing-Open-Source-Digital-Library-Software.pdf

Appendix A: Platform Full Evaluation

0 = N/A   1 = Poor match   2 = match   3 = good match

	Platforms					
	Omeka	Notes	AtoM	Notes	Islandora	Notes
Required						
Ease of use	2	Evaluated through presentations and own testing of platform.	2	Evaluated from presentations.	1	The majority of articles, evaluations etc were from a developer/IT perspective with no mention of ease of use. Set up is not straightforward - requires quite a bit of time therefore greater cost.
Handling of controlled vocabularies and entitiy authority	2	Can use some controlled vocabularies (via plugins); can set up own controlled vocabs.	3	Provides strict entity authority, taxonomies	2	Entity Solution Pack; you are able to provide controlled vocabs or external links but not straightforward to implement
Metadata standard support and ingest/export facilities	1	Dublin Core but other metadata standards available via plugins. Import via CSV, OAI-PMH. Export via Omeka and METS XML, other formats	3	ICA description standards. RAD, DCMI, MODS. Batch import via CSV; import/export using EAD XML	3	Any metadata schema supported; RDF support for any XML standard. Batch metadata ingest; export via XML

Support (active user and developer communities)	3	Very active user and developer communities. At least some developers available in UK/EU	3	User support extensive; at least a few developers in the UK	3	Active user and developer communities
Support of learning resources and any online learning activities	3	Supported - lesson plans and interactive resources (somewhat)	2	Not explicitly stated but additional files are fine. Interactive resources not mentioned (therefore unlikely)	2	Not explicitly stated but can add files and set up relationship between learning resources and source files (interactive resource support not mentioned)
Handling of external links	3	Hyperlink support	2	Hyperlink support but only in fields that are uncontrolled	2	hyperlink support (couldn't find references on how easy it is to do or where you are able to put them)
Cost[8]	2	OS - configuration and maintenance costs only	2	OS - configuration and maintenance costs only	1	Given the complexity of the bundled platform development costs likely to be pricey (from our perspective)

---

[8] To be a good match (a 3) the software would not require any developer time to implement and configure.

OAI-PMH compliance	3	OAI-PMH repository plugin	3	OAI-PMH repository plugin. Requires authentication	3	Islandora OAI module
Search API	3	REST API	2	Elasticsearch query through REST API	1	Islandora REST API is in development but is focussed on Solr
Full text searching (within PDFs especially)	3	PDFText plugin allows for searching within pdfs. No full-text searching within ePub format	3	PDF full text search available. Not available for ePub.	3	PDF full text search available. Unclear for ePub.
Persistent identifiers for digital objects (URIs)	2	CoolURIs + github plugin (ArkAndNoid4Omeka)	1	Could find no reference within AtoM documentation or in use apart from one example that uses handle.net	2	Can implement external resource identifiers services

Support large format images, video, audio, pdf, epub, word documents and other common formats	3	Large images supported through IIIF support/universal viewer (requires an IIIF server). ePub supported	1	Doesn't seem to support large format images, no zoom function. No IIIF support. No ePub support.	3	IIIF support available. Does support wide range of formats including TIFF. Possibly TEI?
Create exhibitions	3	Exhibitions functionality built in	0	(From forum comment, Jan 2016: At this time, however, we do not know of anyone who is seriously considering sponsoring the development of such a module)[9]	1	Likelihood all exhibitions would need to be built in Drupal - does not appear to be a dedicated exhibition module.
Sum:	**33**		**27**		**27**	

Optional					
Citation export and social media shareability	2	Citation export through Zotero (through COinS); could find no support for Endnote/Bibtex export. Social media sharing plugin available	1	Export to Endnote/Bibtex; does not export to Zotero. Doesn't seem to have social media sharing functionality	3	Citation export to Endnote, Bibtex and Zotero. Social media sharing available
Crowdsourcing functionality (transcription)	3	Scripto (transciption) plugin.	0	Could find no reference to this functionality	3	Crowdsourcing transcription possible through Islandora webform

---

[9] https://groups.google.com/forum/#!topic/ica-atom-users/3O0_yV3wLIU

User contributed material, tags, etc	2	Plugins for user contribution of a variety of types. Commenting allowed but only MyOmeka has tagging. Poster plugin	1	Registered user groups may tag	3	User contributed content possible as is tagging
Total:	**40**		**29**		**36**	