Making African Connections Infrastructure Report

James Baker, January 2019

Scope

As part of the process of documenting the project outputs, an infrastructure report has been written. This report is intended to create a history of the digital archive's construction, recording how and why various technologies were selected. The report includes technical descriptions and rationales for decision related to best practice in areas such as software development (e.g. use of open source software and publishing software in open source repositories). The infrastructure report aims to make the project findings useful for other researchers working on similar initiatives, as well as ensuring best practice for the project itself by providing a reflective mechanism for monitoring the archive's technical development, especially in areas such as code commenting and functionality description.

Overview

The AHRC-funded project "Making African Connections: Decolonial Futures for Colonial Collections" launched in January 2019 with the aim of furthering both conceptual and applied debates over 'decolonizing' public institutions. A Making African Connections Digital Archive was planned as part of the project activity, the aim of which was two-fold:

- To provide public access to historic African collections held by three museums - Brighton Museum & Art Gallery, Royal Engineers Museum, the Powell-Cotton Museum - digitised during the course of the project.
- To enable the project team to experiment with ways of presenting these collections in line with the 'decolonizing' ambitions of the project.

To support the creation of the Digital Archive, the project team used various technologies. This report - produced at the mid-point of the project - describes those technologies, why there were chosen, and how they were implemented. We believe that the products of our labour are hard to understand unless we document that labour. Making ourselves transparent is vital to future interpretations of the historic African collections we are working with.

Infrastructure

Reclaim Hosting

We use Reclaim Hosting for our web hosting and domain registration, registered using a shared email address (makingafricanconnections{at}gmail[dot]com). We chose Reclaim Hosting for the following reasons:

- It is geared towards the educational community, with a number of education focused web applications available via its Installatron plugin.
- The Installatron plugin reduces the technical skills required for server management, whilst - at the same time - enabling bespoke control when required.
- It has built in tools to measure site traffic, create server backups and recovery points, and manage data security.

Risks associated with the hosting are as follows:

- Although login details for the Reclaim Hosting Client Area and the server cPanel are shared, within the team only Baker has the skills/knowledge required to configure the server and resolve issues (e.g. rolling back the site to a recovery point after a failed software install).
- Due to the high-quality nature of images used in the project, we currently pay $100 annually for 100GB of storage. Whilst this can be reduced to the 10GB plan ($50 per year) after the batch uploading phase of the project is complete, no budget is in place for sustaining the resource after AHRC funding has expired.
- The carbon output of Reclaim Hosting servers was explored with the vendor, however their provider - Digital Ocean - does not track energy usage on specific servers (only general information). As Sussex Humanities Lab transitions to less carbon intensive and ecologically destructive infrastructures, this will need to be monitored.

Omeka S

Omeka S was chosen as a publishing platform for the Making African Connections Digital Archive. It was chosen because it is open source, built for cultural organisations to share collections, community supported, built on well-used technology standards, allows for flexible representation of things and concepts, provides granular access controls, and has an GUI user interface comparable to commonly used software like Wordpress. As Baker could manage to deployment and maintenance of Omeka S, our development budget was spent on customisations that optimise mobile use in low resource environments.

Risks associated with this publishing platform are as follows:

- After enquiries, it emerged that the Omeka team have conducted only limited research on user experience of Omeka sites (their focus has been on user experience for creators of Omeka sites). It is unknown, therefore, how users will respond to Making African Connections.
- Customisations were made before the release of Omeka S v2 and our custom 'Theme' is not compatible with Omeka S v2. Therefore, updates to Omeka S have been disabled, which may impact on the sustainability of the service.

Wordpress

Wordpress was chosen as the publishing platform for the Making African Connections Digital Archive. It is was chosen because it is a widely used publishing platform which the project team had familiarity with.

Storage and preservation

We use a combination of flash drives, cloud storage (University of Sussex implementation of Box), and SHL's secure Network Attached Storage Device for in-project storage. Our digitisation workflow sees the data we create move from the least secure (flash storage) to most secure (NAS device) as the data is refined and prepared for preservation in the University of Sussex Research Data Repository. This latter service is built on Figshare and links to Arkivum, and therefore is geared towards long-term preservation of project assets.

Shell tools

We use a variety of open source shell tools to enable quick, scaleable, and reproducible processing of data for a variety of tasks. These are hosted on our Github repository, and include the following processes:

- Bulk renaming of image files to establish links with associated objects and documentation.
- Bulk conversion of image files from preservation formats (.CR2, .tif) to display formats (.jpg) using Imagemagick
- Bulk generation of preservation metadata and file verification using Siegfried.

Dublin Core

We use Dublin Core as a metadata schema for the federation of documentation produced by three museum collections, each of which as based on different museum cataloguing systems. Dublin Core was chosen for this purpose for the following reasons:

- It is widely used standard.
- All metadata elements are optional and repeatable, meaning that we can create space for multi-vocality, uncertainty, and dispute in our metadata.
- It can be used to represent non-collection items (such as people and materials) as well as collection items.
- It is supported by Omeka S.

GitHub

We use the code repository GitHub to host documentation, code, and workflows relating to the project. Based on versioning software, GitHub creates a public and transparent record of changes to our project processes that wil exist beyond the life of the project.