E X A M P L E – Data Management Plan

Source: https://www.dcc.ac.uk/resources/data-management-plans/guidance-examples

This is a fictional AHRC technical plan developed by a reviewer to highlight common pitfalls.

Technical Plan

Section 1: Summary of Digital Outputs and Digital Technologies

The digital output of the project will be a database of historical recipes from throughout Britain from the 1600s to the first world war. Recipes will be geocoded based on project research into where the recipe originated and it will be possible to search for ingredients or styles of cooking based on region / town in combination with dates to show how the use of ingredients and methods changes over time and space. A highly innovative, vibrant and interactive public website will be created through which data will be plotted on historical map layers and on interactive visualisations, and an app will also be released. User interaction will be integral to the website and users will be able to post comments about recipes including images and video clips of dishes created by following the recipes. A content management system will also be developed for the project's researchers to use to record recipe and other associated data over the course of the project and for user submitted content to be managed.

Section 2: Technical Methodology

2a: Standards and Formats

The project will use a variety of open and proprietary formats that will best suit the needs of the project's outcomes. These will be migrated to suitable open standards to facilitate preservation at the end of the project. Text will be transcribed as plain text with HTML markup and will take up roughly 500Mb of space. Images will be in the JPEG format and 5Gb of server space will be set aside for them. Video files will be MOV and 20Gb of space will be available for them. Visualisations will be SVG files. The map interface will be based around Google Maps. Web pages will follow current HTML and CSS standards.

2b: Hardware and Software

The website and content management system will be hosted on LAMP servers based at project partner the University of Edinburgh who will supply the project with two virtual servers: a 'development' and a 'live' server. The resource will be developed using the Joomla content management framework. Images will be edited with Adobe Photoshop and videos with Final Cut Pro.

2c: Data Acquisition, Processing, Analysis and Use

An initial project website will be set up by the developer during the first month of the project. This version of the website, along with project presences on social networking sites such as Facebook and Twitter, will be managed by the Co-I. During months 1-3 the developer, in collaboration with the rest of the project team, will create a scoping study for the content management system and the public website. S/he will work on a first version of the CMS during months 4-6, launching it during month 6 with further iterations which will introduce further functionality being made every few months over the three years of the project.

The 10 project RAs, who will have begun collating data on their laptops from fieldwork to libraries and archives around the UK in month 2, will receive training in the use of the CMS by the developer

in month 6. Data uploaded to the CMS by the RAs will be analysed and processed by the developer to convert it into formats suitable for display on the project website, which the developer will be working on during years 2 and 3 of the project.

A 'beta' version of the online resource will be made available to selected users midway through year 2 of the project. This version will feature full access to the recipe records but limited search and browse functionality. The launch of the 'beta' version will coincide with the project symposium.

The developer will continue to refine the online resource, adding functionality such as the map interface and the visualisations throughout year 3 of the project. An official launch of the final resource that will be available to all users will take place in the final month of the project to coincide with the project conference, at which point users will be able to access the maps and visualisations and post comments on the recipes.

The online resource, the CMS, one IIIF server, and the Solr indexing system will be located on Linux webservers managed by University of Glasgow IT Services. The webservers will be backed up nightly to an Ultrium LTO2 unit located remotely from the server. The project archive will be stored on an Active Directory network, supported by 4 domain controllers, located in two separate 'server' rooms at each end of campus. Each server has a RAID disk subsystem and is backed up nightly to devolved backup systems. Both server rooms are protected by both UPS and generators. The backup system creates and maintains two copies of each system state backup which are held on near-line disk, on-site tape and off-site tape. 7 versions of each AD state are retained for 90 days. The data schema, system specification and procedures for data creation and management will be described in a detailed set of documents.

Section 3: Technical Support and Relevant Experience

The PI and Co-I both have considerable previous technical experience. The project aims to recruit the developer before the start of the project but at the very latest s/he will start working by the end of month 1. The developer will be based at the University of Edinburgh and will work closely with other project partners at Edinburgh and the University of Glasgow. The PI and Co-I have consulted widely with several other projects of similar aims and complexity in order to gain a better understanding of how technology can be best harnessed in order to make a truly groundbreaking digital resource. Advice on data management has been sought from the AHDS.

Section 4: Preservation, Sustainability and Use

4a: Preserving Your Data

Upon completion of the project the digital outputs of the project will be migrated to open standards for preservation as discussed in Section 2a. Outputs will be tagged with appropriate metadata to facilitate their discoverability. Long-term preservation of digital data is a considerable challenge; however, how best to preserve digital data is not the focus of this project and other projects within our partner institutions are already making significant progress in how this issue can be addressed.

4b: Ensuring Continued Access and Use of Your Digital Outputs

The website will continue to be hosted by the University of Edinburgh beyond the end of the project and user comments and contributions will continue to be enabled. In order to minimise Brian Aitken, 2015 the cost of sustaining the resource in the longer term the CMS will not be retained following the end of the project and the focus instead will be on the public facing website.