Data Curation Profile for History

Researcher: Graham J. Black, PhD candidate, School of History, Classics & Archaeology, University of Edinburgh

Project: Aerial bombing during the Vietnam War

Project dates: 2012-2015

Profile author: this profile was compiled by Margaret Forrest, Academic Liaison Librarian, School of History, Classics & Archaeology, following an interview with Graham Black on 23rd July 2014.

1. OVERVIEW

1.1. Overview of the research project

The purpose of this research project is to analyse the aerial bombing of Vietnam by US military forces during the Vietnam War and especially during the period from 1965 to 1968. The research proposes to demonstrate how media reporting, the print press in particular, helped to define the public image of this campaign and how this contrasted with the official line taken by the U.S. administration.

1.2. Funding source

Wolfson Foundation

1.3. Data management planning for project

There was no data management plan established for this project at the outset and this is regretted by the researcher. No plan was required by the funders and the researcher was not aware of the need for this at the start of the project.

1.4. Overview of the data and research records related to the project

This comprises mainly US Government documents relating to the bombing, contemporary articles in the press (e.g. New York Times, Washington Post) and the researcher's abstracts and notes from study of the primary sources.

1.5. Intellectual property owner of the data

For the three different types of data mentioned above this would be the US Government, newspaper publishers or journalists and the researcher, respectively.

1.6. Approximate number of data files generated during the course of the project

24,000 files

1.7. The average size of the data files

Average per single downloaded file 65.5 Kb for newspaper articles and 150 Kb for photographic images of documents. 250-300 Mb within collective folders. Approximate total of 6,000 Mb.

2. ORGANISATION

2.1. Data formats

The files are in .docx, .pdf and .jpeg format.

2.2. Organisation of data

The files are structured using the date of download, by month and year.

2.3. Importance of the metadata system used for project

The researcher would like to retain his current system of organising data, but would be willing to try another system to complement this, especially if it would help save time and effort in retrieving information from files.

2.4. Software programs or tools used in the collection and organization of this data

Data was collected by downloading .pdf files from primary source databases (e.g. newspaper archives) and by photographing print only documents. No software programs were used in the organisation of this data other than that available on the computer's folder system.

2.5. Software programs or tools to utilise this data

Adobe Reader, MS Word, Windows Photo Viewer

2.6. Storage and back-up of files

Files are stored on the researcher's University home drive and on USB memory sticks. The USB devices were initially used only for back-up. However, more recently they have been used as the only storage place for some data, due to the fact that the researcher had reached his limit of available space on the University home drive. The researcher was not aware of the research data storage facility offered by Information Services before the day of the interview.

2.7. Measures to control access to data

EASE login password; USB devices are not currently encrypted.

3. STORAGE & SHARING

3.1. Future preservation of data

The researcher would like data from the project to be preserved indefinitely.

3.2. Publication of research and data linking

The researcher very much hopes to publish the results of this project and make openly available primary sources where permission is granted by the copyright holders.

3.3. Intended audience of the data

The data has a wide ranging potential audience and capacity for further analysis in fields such as American and Vietnamese history, sociology, politics and the study of modern conflict.

3.4. Reasons for the choices made in the 'data sharing matrix'

Some choices in the sharing of data were restricted mainly for two reasons. Firstly, due to copyright legislation, e.g. permission would first need to be obtained from copyright holders of articles in newspaper archives. Secondly, the sharing of some data could put

individuals or their families at risk, e.g. Vietnamese informers to the US Government during the War

3.5. Conditions or constraints placed on the sharing of this project data

None; the only requirement by the funders was that the organisation be cited at the point of publication.

3.6. Requirement for data sharing usage statistics and measurements of use

The researcher would be interested to know how many people use this project data in future; how they use it; where and how often it is cited.

3.7. Data sharing embargoes

The researcher would prefer if there could be no embargo or special permission required to access this data, but this would depend on the copyright and "sensitive data" restrictions already mentioned.

3.8. Anticipated future use of data

The data could be used to enable future researchers and historians to come to their own conclusions on the subject. It is hoped that lessons could be learned from this study of modern conflict.

3.9. Additional support services desired at the University of Edinburgh

The researcher was unaware of the support available for the planning and storage of research data at the University and suggested that this could be better highlighted to all relevant staff and students.

Data-sharing matrix: data types and levels of sharing anticipated

List each type of data here (planning documents, raw data, analysis, etc)	Wouldn't share with anyone	Would share only with my collabor-ators	Would share with others in my field	Would share with other academics outside my field	Would share with the general public
.pdf files from newspaper archives (under copyright)	✓				
Photographs made by researcher of print only restricted primary sources, e.g. US Government documents	✓				
Maps drawn by researcher		✓	✓	✓	✓
Abstracts and interpretations of primary sources written by researcher		✓	✓	✓	✓
Audio files of interviews where permission to share is not granted (e.g. aircraft pilots)	✓				