**Understanding and Improving Research Data Management in the Visual Arts:**

**Case Study of the KAPTUR Project**

Dr Robin Burgess. The Glasgow School of Art, 167 Renfrew Street, Glasgow, G3 6RQ, United Kingdom. (r.burgess@gsa.ac.uk)

*“What is arts research data? What does it mean to you? Research, art, design, architecture, I’m going to tell you. What is arts research data? We tried to find out. We asked various researchers, and this is what we found…”*

Research data is a valuable resource and, with appropriate curation and management, it has much to offer learning, teaching, research, knowledge transfer and consultancy in the visual arts. However, very little is known about the curation and management of this data: none of the specialist arts institutions have research data management policies or infrastructure in place and evidence suggests that practice is ad hoc, left to individual researchers and teams with little support or guidance. In addition the curation and management of such diverse and complex digital and physical resources is challenging.

Led by the Visual Arts Data Service, a Research Centre of the University for the Creative Arts, in collaboration with the Glasgow School of Art; Goldsmiths, University of London; and University of the Arts London, and supported by JISC, the KAPTUR project (2011-2013) sought to address this lack of awareness and explore the potential of appropriate research data management in the visual arts, by:

* Uncovering the nature of research data in the visual arts
* Investigating the current state of management of research data
* Developing a model of best practice applicable to specialist arts institutions and arts departments in multidisciplinary institutions
* Applying, testing and piloting a model for best practice with the institutional partners

This paper shares the experiences that the Glasgow School of Art encountered while being part of the KAPTUR project.

Across the higher education sector, research councils, organisations, teams and researchers are under pressure to make publicly funded research data freely available, and in line with the research councils UK guidance, the publication of data is increasingly a requirement of funding. Equally important is data transparency and the ability for researchers to access data to test the validity and reliability of the research outputs and methods; to reinterpret and reuse data, thereby adding value to publicly funded research; and ultimately being able to access the data in the longer term.

Several universities, for example the University of Edinburgh, have established institutional research data management policies and infrastructures to support the deposit, curation and preservation of research data, whilst others are in the process of piloting services. In addition, several of the Research Councils support repositories themselves, for example the UK Data Archive at the University of Essex which is funded by ESRC. JISC have also funded several projects and research strands in the field of research data and its management, to provide for addressing the issues involved in the curation, management, preservation and usage of research data.

However, until recently, little work has been undertaken to understand the distinctive and varied nature of research data in the visual arts, and even less to understand how this data can be curated, managed and preserved appropriately within the higher education sector.

By its very nature, research in the arts is highly complex and varied, often comprising a wide variety of outputs and formats which present researchers, information managers and technology teams with many discipline specific issues. The methods and processes which generate this research information are just as varied and complex. Research in the arts relies heavily on sketchbooks, logbooks, journals and workbooks. The often physical nature of research in the arts and its security and preservation also presents researchers and curators with significant problems and greatly increases the risk of data loss and deterioration. Alongside this data, a wide range of related research documentation and protocols are also created.

However the appropriate curation and management of research data in the visual arts is extremely desirable for a number of reasons:

* It satisfies funding requirements and demands for open access
* It helps to demystify research methods and outputs, and the arts themselves
* The production of research data is extremely time consuming and therefore costly and its lack of discoverability or loss is inefficient
* It may have a significant application and value far beyond the cost of its initial collection and outcomes which could have gone unrecognised in the initial investigation
* It would enable other researchers to test the reliability and validity of the data and research method
* Open access to research data would enable its reuse and fascilitate new interpretations, thereby adding value to and enhancing the impact of the original data
* The publication of data would enable it impact to be tracked more accurately
* It would enable researchers and semantic web tools to make new links between isolated and previously undiscoverable datasets more easily
* It would extend collaborative opportunities between researchers and teams working on similar and related datasets to create new research opportunities

As such, the appropriate curation, management and preservation of research data in the visual arts can be considered not only as desirable but as essential.

It can be stated that the objective of the KAPTUR project was therefore to investigate the nature of research data in the visual arts and then to consider the application of technology to support collection, discoverability, usage, and preservation of research data in the arts. To help this process of discovery, reviews were undertaken in relation to policies, procedures and systems, and case studies were developed to showcase good practice to a wider audience.

In the context of GSA, our involvement in KAPTUR has been of great importance as prior to embarking on the project there were no research data management policies in place, few considerations towards the required infrastructure for sound research data management, and limited technology and systems in place for the management of research data.

The first piece of analysis undertaken to help understand the nature of research data management at GSA was the ‘environmental assessment’. This was followed by the technical and systems infrastructure assessment.

The environmental assessment was undertaken by all partner institutes by carrying out interviews with researchers where questions were asked in relation to:

1. Terminology: Questions were asked to determine how much the researchers understood about research data management in relation to methods and processes. Also, whether they were able to define the concept of research data in their minds.
2. The role of the visual arts researcher: We wanted to try and understand how researchers work with the data they collect, the interactions they make. Their role was looked at from an individual perspective, but also from an institutional perspective.
3. Creation of visual arts research data: How is arts research data created? What are the processes that researchers adopt when creating their outputs? What contributes to their research outputs? We were keen to understand if any specific method are used, any common ground covered in creating the data, however for visual arts this can be quite varied.
4. Use/re-use of visual arts research data: With all research it is important to be able to use and re-purpose information if required. The views of researchers were gathered in relation to this concept. This was to help determine the researchers views towards the ability to re-use the work/data they have been creating.
5. Visual arts research data in the longer term: As with all data, plans need to be put in place for the storage and accessibility of the information, and whether it needs to be kept for a period of time etc. Storage methods were discussed and the thoughts of the researcher on the preservation of information examined.

The outcome from the report was focused on understanding the state of play for research data management at the partner institutes, particularly from the view point of the researchers. It also helped to develop further questions and triggers in relation to the requirements that the institutes needed. An additional outcome from the environmental assessment was the development of an A-Z of arts research data, that looked at assigning a word and a phrase to each letter of the alphabet and presenting this in a poster that could be used as a quick reference guide for researchers and research/information services staff when tackling research data in the arts. The following image illustrates the terms and images determined for the word ‘KAPTUR’.



Fig 1: A-Z of arts research data, Kaptur (Designed by Gii Bear, 2012, images courtesy of GSA Flickr)

With regard to the technical development within the KAPTUR project, two elements were of interest.

(1) To determine the user and technological requirements to support the effective management of research data in the visual arts, and

(2) To identify and recommend a technical solution to support the effective management of research data in the visual arts.

Questions were asked of each institute about the types of data they collect and manage, disk space requirements, shared drives and remote servers, operating environments, the Cloud, authentication methods, tracking of research data, back up procedures, and required support for the technical aspects of data management. In answering these questions a report was written that stipulated the technical requirements for the management of research data. Requirements covered included: storage, interface, system and institutional.

With regard to what system to apply for the management of research data, this is still under discussion amongst the project partners. Options reviewed have included Dataflow, DSpace, Fedora, Figshare, CKAN and EPrints to name a few. As all partner institutes have research repositories developed using EPrints, there has been great focus on this software as a solution. I new module has been developed and tested that allows research staff to upload information related to their research data, procedures and processes, not just the actual research output. This information can be linked with the final output that would be viewed within the external facing research repository. The process for inputting research data to the system is similar to that of depositing a research output within the repository. This helps to eliminate the need to train staff in using the potential management system, and should help to reduce issues with the software/system. The strengths that the use of EPrints has brought to the management of research data include:

* EPrints can accommodate different workflows; these can be edited to provide different options
* Content can be stored in any file format as designated by the repository administrator during configuration. Multiple representations of the same content are also permitted
* With the release of EPrints version 3.3 (September 2011) repository managers can install applications with ‘1-click’ through the EPrints Bazaar, described as an ‘App Store’. These applications can be downloaded and installed in the repository without affecting the core configuration and original settings of the repository. The application can also be easily disabled or deleted.

As well as strengths there are also a few weaknesses in applying EPrints as the software solution. These include:

* EPrints is open source software and relies of project funding to ensure upgrades can be supported
* A series of plugins may need to be installed and tested before setting up the production environment
* Some configurations may need to be performed manually.

This has been the approach undertaken for the KAPTUR project, the potential application of EPrints for helping to improve the management of visual arts research data through the development of an additional module that can feed into existing institutional repositories. It must however be noted that there is no single product that can fully capture all the requirements identified within the KAPTUR project, EPrints was the best fit.

It can be concluded that research data in the visual arts can be seen as complex and complicated, it does not always fit into the natural scheme of data management. Therefore the development of policies for arts related data management and systems/infrastructure should be aware of thinking outside the box, and lending themselves to being more iterative and open to interpretation. Many research data management policies take the approach of being ‘aspirational’ in nature in that they elude to processes and requirements, rather than being completely set in stone. They look to build on the methods already in place at the institute, and subsequently as plans and strategies develop, the policy and infrastructure can also be enhanced and modified. This is very much the approach that GSA has taken for the development of their policy and data management infrastructure.