From KAPTUR to VADS4R: Exploring Research Data Management in the Visual Arts

*Robin Burgess**

Across the higher education sector, research councils, organizations, teams, and researchers are under pressure to make publicly funded research data freely available, and in line with the Research Councils UK guidance. Publication of data resulting from the research is increasingly a requirement of funding. Equally important is data transparency and the ability for researchers to access data in order 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, to access the data in the longer term.

By its very nature, research in the visual arts is highly complex and varied, often comprising a wide variety of outputs and formats that present researchers, information managers, and technology teams with many discipline-specific issues. Examples include sketch books, paintings, architectural plans and buildings, physical artifacts, and complex modelling algorithms. Additionally, the methods and processes that generate this type of research information are just as varied and complex. Research in the visual arts relies heavily on sketchbooks, logbooks, journals, and workbooks. Alongside this data, a wide range of related research documentation and protocols (such as "how-to guides" and methodology reports) are also created. The physical nature of research in the arts presents researchers and curators with significant problems with security and preservation issues while also greatly increasing the risk of data loss and deterioration. Issues arise, for example, in the field of architecture. When data is locked up in the physical building that has been created as the output, how can this information be preserved and managed?

Therefore, appropriate curation and management of research data in the visual arts is essential to
  - satisfy funding requirements and demands for open access;
  - limit the issue that production of data can be extremely time-consuming and therefore costly;
  - reduce its lack of discoverability or loss;
  - enable other researchers to test the reliability and validity of the data and the research method;

- enable greater impact of research and make tracking more accurate; and
- extend collaborative opportunities between researchers and teams working on similar and related data sets to create new research opportunities.

INTRODUCING THE KAPTUR PROJECT

Led by the Visual Arts Data Service (VADS), a Research Centre of the University for the Creative Arts (UCA), 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 these problems. Each partner institute had been engaged in projects together such as KULTUR and shared a common need for managing research data in the arts and therefore joined forces for the KAPTUR project. A steering group was set up to help manage the continuation of the project, and each institute provided a project officer to carry out the required work for the project. Regular meetings were held between the partners, alternating between locations—Glasgow and London. A project website was set up to help capture the work that was being undertaken and to share this with the wider community.

Visual arts research data is a valuable resource and, with appropriate curation and management, it has much to offer in learning, teaching, research, knowledge transfer, and consultancy in the visual arts. From the outset in 2011, the KAPTUR project team noted that very little was known about the curation and management of this data: none of the specialist arts institutions (e.g., the partners involved in the project, like Glasgow School of Art, and other arts-based institutes) had research data management policies or infrastructure in place, and evidence suggested that curation practice was ad hoc and left to individual researchers with little support or guidance. In addition, the curation and management of such diverse and complex digital and physical resources presented unique challenges. These challenges were associated both with the curation (management, handling, storing the data) and the preservation of the data for reuse—such as collecting numerical data and algorithms that constitute the modeling of the human body as demonstrated by the work done by the Digital Design Studio at The Glasgow School of Art, or capturing the processes and ephemeral thoughts created during a dance performance.

The objectives of the KAPTUR project were two-fold: first, to investigate the role of research data in the visual arts; and second, to consider the application of technology to support collection, discovery, use and reuse, and preservation of research data in the arts. To support this, a number of policies, procedures, and systems were reviewed and case studies were developed to assess and understand

existing tools, knowledge, and practices regarding research data management more generally.

The project began with an environmental assessment that considered issues of terminology, the role of the visual arts researcher (within the institute and externally), and how visual arts research data is created, used, and preserved (the researchers role in this process). Next, a technical review considered two questions: first, what did researchers need to support effective research data management in the visual arts? and second, what was the most appropriate technology solution to facilitate the appropriate management of research data in the visual arts? Regarding the first question, each KAPTUR partner considered the types of data that it collects and manages, disk space requirements, where the data was stored (such as shared drives and remote servers), operating environments, the cloud (whether this would be a reliable option for storage and access to data, and its sustainability as a resource), authentication methods, tracking of research data (in relation to its production and ultimate use within the research project workflows), backup procedures, and required support for the technical aspects of data management. Once the requirements were identified, a variety of potential technical solutions were identified for piloting and review. These solutions presented themselves via the repository platforms that were available to the project partners, such as ePrints and DSpace, but also alternative data management tools, including figShare and CKAN, that the team had researched.

The technical report highlighting findings and methodology for this process can be accessed here: http://www.vads.ac.uk/kaptur/outputs/Kaptur_technical_analysis.pdf.

Finally the team developed two training sessions (which were part of the original project scope, to help disseminate findings), each one hour in duration. The first looked at the basic principles of research data management in the visual arts, and the second focused on the creation of a research data management plan. These were further developed by each partner into a pilot course using appropriate content and resources from their respective institutions with reference to their particular organizational practices, processes, and disciplinary areas. Each partner worked with different stakeholders (to help develop the training, but also to take part in it). Participants included early-career researchers, research students, established researchers, and professional support staff from a range of departmental perspectives (administrators, IT, library, etc.). Feedback from attendees was favorable, and evaluation results indicated that participants' understanding of research data management had improved or improved considerably. The criteria for assessing this were solely reliant on the perception of those attending the training and how they felt the sessions went and what they had learned. These training materials were then published freely to the higher education community.

Details of the training sessions and the findings include
- Training plans: http://www.slideshare.net/kaptur_mrd/kaptur-rdm-trainingplan
- GSAs institutional workshop: http://www.slideshare.net/kaptur_mrd/tag/gsardmtraining
- Outputs and Toolkits: http://www.vads.ac.uk/kaptur/outputs/index.html

VISUAL ARTS DATA SKILLS FOR RESEARCHERS (VADS4R) PROJECT

Following the work of the KAPTUR project, the Visual Arts Data Skills for Researchers (VADS4R) project extended this work by tailoring these learning materials for use with early-career researchers and postgraduate students in the visual arts to inform, support, and embed appropriate research data management practice across the visual arts. Led by the Centre for Digital Scholarship (formerly known as VADS) at the University for the Creative Arts (UCA) and in collaboration with Falmouth University and Glasgow School of Art, VADS4R has developed, delivered, and evaluated training programs at each partner institution. VADS4R ran from February 2013 to July 2014 and was funded by the Arts and Humanities Research Council (AHRC). The approach, method, and lessons learned from the project can be accessed from the project website, http://www.vads4r.vads.ac.uk/p/welcome.html. Here you can access the training packages that were developed (using XERTE software) as a significant output from the project. These looked at:
- How to avoid a data disaster
- Writing an AHRC technical plan
- Introduction to research data
- Data management planning
- Managing the material
- Principles of research data management in the visual arts
- The discoverability and reuse of visual arts data

These training packages can be found at http://www.vads4r.vads.ac.uk/p/online-learning.html. Also here you can see the additional material used as part of the training programs developed and the findings determined during some of the sessions run at the partnering institutes. Research in this area of data management in the visual arts continues between UCA and GSA, and new methods and approaches will be shared.

In conclusion, research data in the visual arts can be characterized as tangible and intangible, digital and physical, heterogeneous and infinite, and complex and complicated. It does not always fit into the natural scheme of data manage-

ment. However, the development of policies, procedures, systems, and training can provide an innovative and flexible approach for this data, which is iterative and open to interpretation. These approaches support appropriate curation and management of data to alleviate the issues surrounding funder requirements, elements of time and discoverability, and at the same time improve the impact of research and create new collaborative opportunities for the institutes.

Summary of Step 7.0: Preservation of Data for the Long Term

7.1 Plan for Long-Term Reuse: The delivery and use of the data will rely on long-term preservation planning that anticipates format obsolescence and storage failures.

7.2 Monitor Preservation Needs and Take Action: Actively monitor the integrity and reusability of the data files using appropriate software, and apply digital preservation strategies.

Notes

1. Daniel W. Noonan, "Digital Preservation Policy Framework: A Case Study," *Educause Review*. July 28, 2014, http://er.educause.edu/articles/2014/7/digital-preservation-policy-framework-a-case-study; Nancy McGovern, "Digital Preservation Management Model Document," Version 3.0, Digital Curation and Preservation Framework: Outline, last revised September 2014, http://www.dpworkshop.org/workshops/management-tools/policy-framework/model-document.

2. The pdf version and the web version are available online: University of Minnesota Libraries, "Digital Preservation Framework," last updated January 2014, https://www.lib.umn.edu/dp/digital-preservation-framework.

3. Read about sustainability factors used by the Library of Congress's Digital Formats website: Library of Congress, "Sustainability Factors," Sustainability of Digital Formats, Planning for Library of Congress Collections, accessed March 15, 2016, http://www.digitalpreservation.gov/formats/sustain/sustain.shtml.

4. Library of Congress, "Recommended Formats Statement, 2016–2017," accessed August 7, 2016, http://www.loc.gov/preservation/resources/rfs/RFS%202016-2017.pdf.

5. For environmental storage conditions for magnetic tape, CD-ROM, and DVD, see Ross Harvey, *Preserving Digital Materials* (Berlin: Walter de Gruyter, 2005), 125–26.

6. Note that replication is not the same as backup. In a typical storage backup system, a copy of the data is replicated at multiple locations. When a change is made to one copy, that change gets mirrored to the other locations according to the backup schedule (e.g., nightly). Therefore, if one copy is compromised, perhaps through bitrot deterioration, then all copies will be replaced with this corrupt version.

7. Such as those found in the BitCurator Environment (BitCurator homepage, accessed