Visual Resources as Research Data Management & Digital Curation

Hannah Marie Marshall
ARLIS/NA 2017
Scope Drift: New Directions in Visual Resources
Visual Resources as Research Data

1. Kultur (JISC)
2. Kultivate (JISC)
3. CAiRO (JISC MRDP)
4. KAPTUR (JISC MRDP) - Environmental Assessment Report

http://bit.ly/2kIW1gp
Defining Visual Arts Research Data

“Research data can be described as data which arises out of, and evidences, research. This can be classified as observational, including: sensor data; experimental; simulation; derived or compiled data for example databases and 3D models; or reference or canonical for example, a collection of smaller datasets gathered together. **Examples of visual arts research data may include sketchbooks, log books, sets of images, video recordings, trials, prototypes, ceramic glaze recipes, found objects, and correspondence.”**

**Research Output** - regardless of presentation, is a planned public statement of new knowledge or interpretation

**Data set** - research collated in a certain way to substantiate a particular interpretation, analysis, or argument. May not always lead to a research output (Garrett and Grandstadt 2012)

**Evidence** - Evidence which is used or created to generate new knowledge and interpretations (Garrett 2012)

http://www.dcc.ac.uk/blog/tackling-visual-arts-research-data
Data Management Planning

- **Data types and sources**
  - What types of data, samples, physical collections, code, software, curriculum materials and other materials will be produced in the course of the project?

- **Formats and standards**
  - What standards will be used for your files and metadata?

- **Roles and responsibilities**
  - What are the roles and who has responsibilities for managing data?

- **Dissemination methods**
  - What are the methods for sharing data and metadata during and after the award period?

- **Policies for public access, data sharing, and re-use**
  - How will you meet funder requirements to provide public access to your data while protecting privacy, confidentiality, security and intellectual property rights?

- **Preservation**
  - How will you preserve the integrity of your data over time?

- **Other considerations**
  - Some funders and/or RFP's require that additional data management areas be addressed in a DMP.

https://data.research.cornell.edu/content/data-management-planning
As part of the KAPTUR project (2011-13), three online toolkits were developed that provide an introduction to research data in the visual arts:

**Toolkits for visual arts researchers**

- Introduction to Research Data
- Data Management Planning
- Managing the Material

**Toolkit for visual arts data managers**

- Managing the Material
More tools

More comprehensive lists of terms are available:
- Digital Curation Centre (DCC) Glossary
- ICPSR's Glossary of Social Science Terms
- OAIS Reference Model
- Society of American Archivists' Glossary of Archival Terminology

http://bit.ly/2kIWr1p

RESEARCH DATA MANAGEMENT SERVICE GROUP
Comprehensive Data Management Planning & Services

Best practices

Best practices documents created by the RDMSG to help you write a data management plan and manage your data.

- Data citation
- File formats
- File management
- Glossary of data management terms
- Guide to writing a Data Management Plan (DMP)
- Guide to writing "readme" style metadata
- Introduction to intellectual property rights in data management
- Metadata and describing data
- Preparing tabular data for description and archiving
- Sharing data
- Support for researchers in the Life Sciences
“Our Curation Lifecycle Model provides a graphical, high-level overview of the stages required for successful curation and preservation of data from initial conceptualisation or receipt through the iterative curation cycle. You can use our model to plan activities within your organisation or consortium to ensure that all of the necessary steps in the curation lifecycle are covered. It is important to note that the model is an ideal. In reality, users of the model may enter at any stage of the lifecycle depending on their current area of need.”

http://www.dcc.ac.uk/resources/curation-lifecycle-model
DCC Digital Curation Lifecycle Model

- Full Lifecycle Actions
  - Description and representation information
  - Preservation planning
  - Community watch and participation
  - Curate and preserve

- Sequential Lifecycle Actions
  - Conceptualize
  - Create or receive
  - Appraise and select
  - Ingest
  - Preservation action
  - Store
  - Access, use, and reuse
  - Transform

- Occasional Lifecycle Actions
  - Dispose
  - Reappraise

http://bit.ly/2kIW1gp
Defining Digital Curation

**Digital Curation** - Digital curation is concerned with actively managing data for as long as it continues to be of scholarly, scientific, research, administrative, and/or personal interest, with the aims of supporting reproducibility, reuse of, and adding value to that data, managing it from its point of creation until it is determined not to be useful, and ensuring its accessibility, preservation, authenticity, and integrity over time ([Oliver and Harvey](https://example.com), 8).

**Digital Preservation** - "all of the actions required to maintain access to digital materials beyond the limits of media failure or technological change" ([Digital Preservation Coalition 2008](https://example.com), 24)

**Digital Archiving** - "the process of backup and ongoing maintenance as opposed to strategies for long-term digital preservation" ([Digital Preservation Coalition 2008](https://example.com), 24)
Scope of Digital Curation

- Refreshing data, checking the accuracy of the refresh, and generating metadata that document the processes applied to refreshing data
- Maintaining the ability to locate digital materials reliably by assigning persistent identifiers to them to ensure they can be found
- Maintaining multiple copies of the bit stream
- Managing intellectual property and other rights
- Recording sufficient representation information for digital objects so that the bit stream is still meaningful and understandable in the future
- Producing digital objects in open, well-supported standard formats
- Limiting the range of preservation formats to be managed
- Keeping track of changes (especially obsolescence) in hardware, software, file formats, and standards that might have an impact on digital preservation, and maintaining multiple copies of the bitstream
Activities of Digital Curation

Maintaining the links between digital information and associated annotations or published materials, including citations

Ensuring the long-term accessibility and re-usability of digital information

Performing archiving activities on digital information such as selection, appraisal, and retention

Ensuring the authenticity, integrity, and provenance of digital information are maintained over time

Performing preservation activities on digital information such as migration or emulation

Maintaining hardware components to enable digital information to be accessed and understood over time

Managing digital information from its point of creation

Managing risks to digital information

Ensuring the destruction of digital information

(Digital Curation, Oliver & Harvey, 8)
Benefits of Digital Curation

Data Creators
- Quality
- Access
- Visibility

Public Good
- Free
- Immediate
- Permanent
- Online

Compliance
- Funding bodies
- Legislation

http://bit.ly/2kIW1gp
Full Lifecycle Actions

**Description and representation information** - Assign administrative, descriptive, technical, structural and preservation metadata, using appropriate standards, to ensure adequate description and control over the long-term. Collect and assign representation information required to understand and render both the digital material and the associated metadata.

**Preservation planning** - Plan for preservation throughout the curation lifecycle of digital material. This would include plans for management and administration of all curation lifecycle actions.

**Community watch and participation** - Maintain a watch on appropriate community activities, and participate in the development of shared standards, tools and suitable software.

**Curate and preserve** - Be aware of, and undertake management and administrative actions planned to promote curation and preservation throughout the curation lifecycle.
Sequential Lifecycle Actions

**Conceptualize** - Plan the creation, capture, and storage of data

**Create or receive** - Create data as well as administrative, structural, and technical metadata

**Appraise and select** - Evaluate data and select for long-term curation and preservation

**Ingest** - Transfer data

**Preservation action** - Ensure long-term preservation and retention of the authoritative nature of data

**Store** - Store the data in a secure manner adhering to relevant standards

**Access, use, and reuse** - Ensure that data is accessible to both designated users and reusers

**Transform** - Create new data by migrating or recombining data
Occasional Lifecycle Actions

Dispose - Dispose of data, which has not been selected for long-term curation and preservation in accordance with documented policies, guidance or legal requirements. Typically data may be transferred to another archive, repository, data centre or other custodian. In some instances data is destroyed. The data’s nature may, for legal reasons, necessitate secure destruction.

Reappraise - Revisit data which fails validation procedures for further appraisal and reselection.

Migrate - Migrate data to a different format. This may be done to accord with the storage environment or to ensure the data’s immunity from hardware or software obsolescence.
Resources


Guidelines for the Preservation of Digital Heritage
http://unesdoc.unesco.org/images/0013/001300/130071e.pdf

VADS4R Project: http://www.vads4r.vads.ac.uk/p/welcome.html

Digital Curation Centre: http://www.dcc.ac.uk/