KAPTUR

ERR, WHAT DO I DO WITH THIS?
EXPLORING INFRASTRUCTURE REQUIREMENTS FOR VISUAL ARTS RESEARCHERS

Leigh Garrett
Visual Arts Data Service
University for the Creative Arts

RDMF10 : 3RD SEPTEMBER : OXFORD

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Err, what do I do with this?

SESSION OUTLINE

- Introductions

- Project Overview

- Method

- Infrastructure Requirements

- Conclusions

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Introductions

VISUAL ARTS DATA SERVICE

- Established in 1996

- National repository for images and metadata: 120,000 items; 300 collections; range of learning and teaching resources; focuses on the visual arts; free for educational use

- Research centre of the university in 2008

- Led and worked on a number of publicly funded projects within the field of digital asset management in the arts

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Introductions

JISC MANAGING RESEARCH DATA PROGRAMME (2011-13)

"Jisc considers it a priority to promote and support good research data management and sharing for the benefit of UK higher education and research."

- 17 institutional projects

- 8 disciplinary projects

- 2 online planning tool projects

- Explore data citation

- Development of training materials

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Introductions

KAPTUR PARTNERS

- Visual Arts Data Service

- Glasgow School of Art

- Goldsmiths, University of London

- University of the Arts London

- University for the Creative Arts

- Jisc

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Overview

BACKGROUND

- Research outputs are varied and complex in the visual arts

- Little is known about the state of research data in the visual arts

- None of the specialist arts institutions have research data management policies or infrastructure

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Overview

OBJECTIVES

- To investigate the nature of research data in the visual arts

- To consider the application of technology to support collection, discoverability, usage and preservation of research data in the area

- To review appropriate policies, procedures and systems within the four partner institutions

- To develop case studies and showcase good practice to the wider higher education sector

Project Overview

STRUCTURE

- Environmental analysis

- Policy formation

- User requirement, systems evaluation and piloting

- Capacity building

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Method

ENVIRONMENTAL ANALYSIS

- Eight informal interviews

- Sixteen in-depth recorded and transcribed interviews

- Literature review

- Data gathered through attendance at meetings and events

- Desk research including reviewing the work of the DCC and previous Jisc managing research data projects

Data gathered informed the user requirement and underpinned the work of the technical work package...

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Method

WHAT IS RESEARCH DATA IN THE VISUAL ARTS?

Tangible
and intangible

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Method

WHAT IS RESEARCH DATA IN THE VISUAL ARTS?

Physical
and digital

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Method

WHAT IS RESEARCH DATA IN THE VISUAL ARTS?
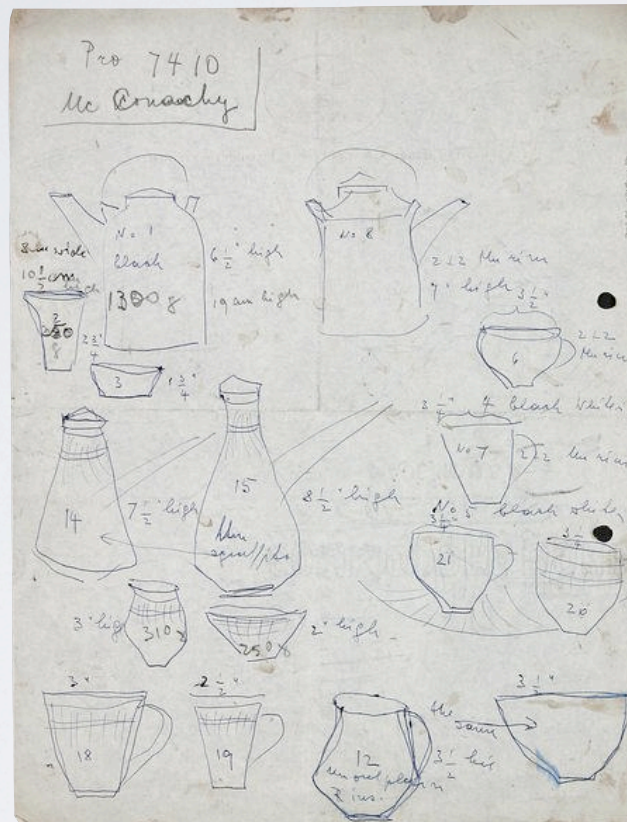
Heterogeneous and infinite

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Method

WHAT IS RESEARCH DATA IN THE VISUAL ARTS?

Complex
and complicated

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Method

WHAT IS RESEARCH DATA IN THE VISUAL ARTS?

"Evidence which is used or created to generate new knowledge and interpretations. 'Evidence' may be intersubjective or subjective; physical or emotional; persistent or ephemeral; personal or public; explicit or tacit; and is consciously referenced by the researcher at some point during the course of their research. As part of the research process, research data may be collated in a structured way to create a dataset to substantiate a particular interpretation, analysis or argument. A dataset may or may not lead to a research output, which regardless of method of presentation, is a planned public statement of new knowledge or interpretation."

Garrett, L. (2013) Defining Research Data in the Visual Arts Online at: http://kaptur.wordpress.com/2013/01/23/what-is-visual-arts-research-data-revisited (retrieved 18 June 2013)

Project Method

POLICY FORMATION

- High level strategy

- Partner institutional working groups established

- Recommendations made to respective research committees

- Approved by all partners

Project Method

USER REQUIREMENT, SYSTEMS EVALUATION AND PILOTING

- User requirement

- Systems evaluation

- Pilot research data management system

- Partner institution adoption

- Err, ongoing...

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Project Method

CAPACITY BUILDING

- Internal dissemination
  - Creation of toolkits and training materials for researchers and professional support staff

- External dissemination
  - Partner institutional case studies
  - Technical case study
  - National and international conferences and papers

- Project conference

Project Method

HIGHLIGHTS

- Definition

- Policy formation

- Infrastructure

- Engagement

Infrastructure

ENVIRONMENTAL ANALYSIS OUTCOMES

- Challenges
  - Collecting research data
  - Managing research data
  - Reuse of research data
  - Preservation

- Researchers agree
  - Research data is important
  - They want to share research data
  - They want to document their research process but do not use any particular standard or methodology

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

RESEARCH QUESTION

Which technical system is most suitable for managing visual arts research data?

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY

The method was initially divided into five stages:

1. Selection

2. Feedback

3. Requirements

4. Technical Evaluation

5. Scoring

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 1

- Selection criteria agreed with partner institutions
  - Solution
  - Storage
  - Interface
  - System
  - Institutional
  - Additional

- Identification of solutions bearing in mind scope and resources of the project

Infrastructure

METHODOLOGY: STAGE 1

- Solution
  - Open source

Infrastructure

METHODOLOGY: STAGE 1

- Storage
  - Metadata storage (25 key fields identified)
  - File storage (39 key types identified)

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 1

- Interface
  - Upload tool for files and metadata
  - Quality assurance workflow
  - Publication of data
  - Preservation of data
  - Data disposal
  - User friendly
  - Search functionality
  - Compliant with W3C standards

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 1

- System
  - Operating system
  - Physical and virtual servers
  - Storage capacity
  - Cloud option
  - Maximum file upload
  - Integration
  - Backup and disaster recovery
  - Security, access and permissions

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 1

- Institutional
  - Workflow
  - Statistical reporting
  - Legal compliance
  - Preservation and data disposal

Infrastructure

METHODOLOGY: STAGE 1

- Additional
  - Mobile access
  - API/Web Service/XML outputs
  - Integration with other platforms such as Eprints
  - SWORD 2 compliant
  - WebDAV interface
  - Able to handle large amounts of data

Infrastructure

METHODOLOGY: STAGE 2

Feedback: partners were asked to select from the requirements the items that they considered to be essential and desirable

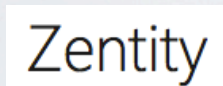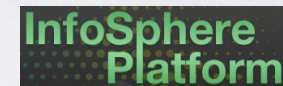| Requirement/Category	Institution A		Institution B		Institution C		Institution D	
	Essential	Desirable	Essential	Desirable	Essential	Desirable	Essential	Desirable
Storage Requirements – capable of handling								
Metadata	√		√		√		√	
Multimedia	√		√		√			√
Text Items	√		√		√		√	
Other types of items	√		√		√			√

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 3

Requirements: seventeen potential solutions were identified

arXiv.org   DataFlow   eprints   Firebird   DSPACE

Dropbox   INGRES   Drizzle   InfoSphere Platform

Zentity   CUBRID   figshare   ckan

MENDELEY   fedora   INVENIO   Google Drive

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 4

Evaluation: each potential solution was scored against each element of the requirements on a simple positive or negative basis

Requirement/Category	CUBRID	DataFlow	Drizzle	DSpace	EPrints	Fedora	Figshare	Firebird
Software Type								
Open Source	X	X	X	X	X	X	X	X
Storage Requirements – capable of handling								
Metadata	X	X	X	X	X	X	X	X
Multimedia	X	Limited multimedia tools	X	Limited multimedia tools	X	Limited multimedia tools	X	X
Text Items		X	X	X	X	X	X	X
Other types of items		X	X	X	X	X	X	X

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 5

Scoring: potential solutions, err...

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 6

Partners graded each requirement numerically (1 the lowest priority and 10 the highest priority) in order to obtain an average score

Requirement/Category	Institution 1	Institution 2	Institution 3	Institution 4	AVERAGE SCORE
Software Type					
Open Source	6	8	6	9	7.25
Storage Requirements – capable of handling					
Metadata	7	7	8	9	7.75
Multimedia	8	7	8	10	8.25
Text Items	8	8	8	10	8.5
Other types of items	9	8	8	9	8.5
Interface Requirements					
Upload tool for files and metadata	9	7	8	10	8.5

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 7

A score was assigned for each requirement against each solution based on: full score if satisfied, part score if partly satisfied and no score if unsatisfied, weighting was equal across requirements

Requirement/Category	Maximum Score	Average Score	DataFlow	DSpace	EPrints	Fedora	Figshare
Software Type							
Open Source	7.25	7.25	7.25	7.25	7.25	7.25	7.25
Storage Requirements – capable of handling							
Metadata	7.75	7.75	7.75	7.75	7.75	7.75	7.75
Multimedia (display)	8.25	5.89	4.125	4.125	8.25	4.125	8.25
Text Items	8.5	8.50	8.5	8.5	8.5	8.5	8.5
Other types of items	8.5	8.50	8.5	8.5	8.5	8.5	8.5

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 8

Finally the scores were totalled and compared

Internal links with other resources such as Eprints systems	6.75	1.45			3.375		
SWORD 2 Compliant	6	3.60		6	6	6	
WebDAV interface	5.5	2.75	5.5	2.75	2.75	2.75	5.5
Able to handle large amounts of data	7.25	7.25	7.25	7.25	7.25	7.25	7.25
TOTAL	**216**	**174**	**171**	**180**	**184**	**159**	**179**

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 9

Five potential solutions were identified and err, ckan

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 10

- Findings published

- Benefits of the systems were considered by the partners within the context of institutional strategic and operating environments

- Piloting of three systems was recommended by the partners and err, ckan

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 11

Piloting figshare with EPrints

- User friendly, visually engaging and intuitive interface

- Flexible file upload tools

- No workflow and quality assurance processes

- Hosted solution

- Poor integration with EPrints for publication of data

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY: STAGE 12

Piloting DataFlow with EPrints

- Open source

- Simple deposit and metadata process

- Poor visual interface

- Developmental and known issues

- Poor integration with EPrints for publication of data

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY REVISED

Err, oops...

- DataFlow: the business and sustainability model

- Figshare: CC0 licensing

- CKAN: retrospective analysis

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

Infrastructure

METHODOLOGY REVISED

Piloting ckan with EPrints

- User friendly, visually engaging and intuitive interface

- Excellent functionality: multimedia, versioning and searching

- Scalable architecture

- Development and stability issues

- Poor integration with EPrints for publication of data

Infrastructure

CONCLUSIONS

- There is no single solution, to completely fulfil all the requirements of researchers, research teams, and their institutions in the visual arts

- In terms of sustainability, licensing and value, ckan appeared to provide the most effective solution

- Time consuming and complex operating environment

- A solution is still required...

Infrastructure

FUTURE?

Cloud

Access Control

Version Control

Interface

Bookmark

EPrints

Harvester

URI and DOI

Citation

Data Uploader

Search

Err, what do I do with this?

REFERENCES

- CKAN, Open Knowledge Foundation (http://www.ckan.org)

- DataFlow, University of Oxford (http://www.dataflow.ox.ac.uk)

- EPrints, University of Southampton (http://www.eprints.org)

- Figshare (http://www.figshare.com)

- Garrett, L., Silva, S., and Gramstadt, M., 2012. Kaptur Technical Analysis Report. Online at: http://vads.ac.uk/kaptur/outputs/Kaptur_technical_analysis.pdf

- Garrett, L. (2013) Defining Research Data in the Visual Arts Online at: http://kaptur.wordpress.com/2013/01/23/what-is-visual-arts-research-data-revisited

Err, what do I do with this?

IMAGE CREDITS

- Pleydell-Bouverie, K. (1930s) pages 20 and 21 of notes on biscuit glazes from notebook. Craft Study Centre. Online at: http://www.vads.ac.uk/large.php?uid=64957

- Rhodes, Z. (1981). The Renaissance/Gold Collection. Zandra Rhodes Digital Study Collection. Online at: http://www.vads.ac.uk/large.php?uid=200470

- Rie, L. Loose sheet with sketches of pots in blue biro. Lucie Rie Archive. Craft Study Centre. Online at: http://www.vads.ac.uk/large.php?uid=22203

- Rie, L. (1910s) Sketches in pencil, ink, and watercolour. Lucie Rie Archive. Craft Study Centre. Online at: http://www.vads.ac.uk/large.php?uid=22392

- Wren, D. (1960s) 21 small tiles used to demonstrate different saltglaze tests on stoneware. Crafts Study Centre. Online at: http://vads.ac.uk/large.php?uid=74581

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC

KAPTUR

ERR, WHAT DO I DO WITH THIS?

EXPLORING INFRASTRUCTURE REQUIREMENTS
FOR VISUAL ARTS RESEARCHERS

Leigh Garrett
Visual Arts Data Service
University for the Creative Arts

VADS : GSA : GOLDSMITHS : UAL : UCA : JISC