DataPool Imaging Case Study: Final Report

**Gareth Beale**, JISC DataPool Project & Archaeology, University of Southampton

**Hembo Pagi**, JISC DataPool Project & Archaeology, University of Southampton

# Executive Summary

The DataPool project’s imaging case study sought to better understand how researchers working with imaging data are coping with the challenges of data management. The project has also been charged with assembling resources and information which will help researchers to develop new approaches to data management as well as making more efficient use of available facilities.  
The study encompassed a wide range of disciplines drawing upon the experiences of researchers from Winchester School of Art, Engineering and the Environment, Physics and Astronomy, Archaeology, Geography, the Library Service, Engineering and the Environment, Ocean and Earth Science, the Centre for Biological Sciences and Electronics and Computer Science. It was divided into three discreet phases:

1.       A Survey of Extant Guidance: The study has assembled a list of data management guidance related to the management of raster and 3D imaging data. As well as supplying a detailed summary of these resources the study has produced a review of the support available to researchers and has produced recommendations as to how awareness and use of these resources might be improved.

2.       Audit of Equipment: An audit of specialist equipment used in the production and use of imaging data has been completed. This data will be added to the list produced by the Southampton-led EPSRC funded Research Facilities and Equipment Sharing Project. It is hoped that this list will encourage efficient and creative collaboration and cooperation.

3.       Data Use Survey: The most substantial aspect of the study has been a survey of a sample of researchers based within the University of Southampton who specialise in the use of raster and 3D data. This survey has sought to refine understandings of existing data management practice. This phase of the study concludes with recommendations as to how data management in these research areas might be enhanced.

## Key Findings

### Survey of Extant Guidance

·         Extensive guidance exists which relates to the effective management of raster and 3D data.

·         Available guidance is unevenly distributed with certain disciplines being much better represented than others. Far more guidance is available relating to the management of image data than for 3D data

·         There is a lack of ‘entry level’ guidance available to researchers unfamiliar with issues of data management.

·         Specialised data management guidance is often difficult to locate

·         Insufficient resources are available which assist researchers in applying general principles to their own work

### Audit of Equipment

·         There has been an increase in the number of specialist devices dedicated to the capture and processing of 3D data since the first list was assembled.

·         Coverage of specialist devices dedicated to the capture and processing of 3D data in the original list of equipment was incomplete.

·        Greater awareness of existing equipment elsewhere in the University would allow more informed decisions to be made relating to on-going investment.

### Data Use Survey

·         Data management strategies are developed at a local level with very little consistency across the University.

·         High standards of data management exist across a wide variety of research groups but these standards are not universal.

·         Greater communication between researchers and research groups would provide a useful means of disseminating good practice.

·         Researchers have encountered difficulties relating to infrastructure, particularly relating to the storage of large volumes of data and the sharing/dissemination of large data sets.

·         Familiarity with data management standards or best practice guidance is variable.

·         Researchers are increasingly reliant upon cloud based media sharing platforms for data curation functionality. This is particularly true where raster data are concerned.

## Recommendations

### Survey of Extant Guidance

·         Guidance should be produced which will help researchers to locate, engage with and utilise appropriate guides to data management. Documents should be easily identified as being relevant to specific subject areas. In addition to this, documents may be produced which introduce the management of specific data types or workflows.

### Audit of Equipment

·        The equipment list should be editable by users and incentives for maintaining the list should be considered.

·        Criteria for including or excluding equipment from the list should be made explicit.

·        Entries in the equipment list should state the availability of the equipment for other researchers and potentially for external work.

### Data Use Survey

·         A forum for sharing and discussing data management strategies would allow good data management planning to attain greater visibility and would encourage best practice across the University of Southampton

·         Guidance documents should be produced which a) lower barriers to participation in good data management practice and b) help users to locate extant guidance which is relevant to their work. These could be supported by a series of workshops in each discipline.

·         Sandbox events could be organised which provide small grants to PhD researchers and researchers to initiate research projects which make innovative use of equipment

# Acknowledgements

The DataPool project is funded by the UK’s Joint Information Systems Committee (JISC)  
Research has been co-ordinated by Gareth Beale and Hembo Pagi both of whom are based in the Archaeological Computing Research Group at the University of Southampton. The work has been supervised by Steve Hitchcock from Electronics and Computer Science and Dorothy Byatt from the University of Southampton Library.   
The project would like to extend its thanks to the researchers mentioned below for their participation, guidance and insights.

* Julian Ball (Library)
* Adrian Baxter (Winchester School of Art)
* Richard Boardman (Engineering Sciences)
* Grant Cox (Archaeology)
* Simon Cox (Engineering and the Environment)
* Steve Crouch (The Software Sustainability Institute)
* Marta Diaz-Guardamino Uribe (Archaeology)
* Richard Dooler (Engineering and the Environment)
* Simon Hettrick (The Software Sustainability Institute)
* Duncan Hornby (Geography)
* Julian Leyland (Geography)
* James Miles (Archaeology)
* Matt Phillips (Library)
* Adam Procter (Winchester School of Art)
* Liam Riddy (Geography)
* Jim Scanlan (Engineering and the Environment)
* Hans Schuppe (Centre for Biological Sciences)
* Tim Sly (Archaeology)
* Mike Street (Engineering and the Environment)
* Fraser Sturt (Archaeology)
* Rex Taylor (Ocean and Earth Science)
* Zondy Webber (Nanomaterials Rapid Prototyping Facility)

Contents

[Executive Summary 1](#_Toc353425296)

[Key Findings 1](#_Toc353425297)

[Survey of Extant Guidance 1](#_Toc353425298)

[Audit of Equipment 2](#_Toc353425299)

[Data Use Survey 2](#_Toc353425300)

[Recommendations 2](#_Toc353425301)

[Survey of Extant Guidance 2](#_Toc353425302)

[Audit of Equipment 2](#_Toc353425303)

[Data Use Survey 2](#_Toc353425304)

[Acknowledgements 3](#_Toc353425305)

[Introduction 5](#_Toc353425306)

[Report Structure 5](#_Toc353425307)

[Aims, Objectives and Outputs 5](#_Toc353425308)

[1. Survey of Extant Guidance 6](#_Toc353425309)

[Methodology 6](#_Toc353425310)

[Introduction to Guidance Documents 7](#_Toc353425311)

[Managing 3D Data 8](#_Toc353425312)

[Managing raster Data 11](#_Toc353425313)

[Findings and Recommendations for Further Work. 13](#_Toc353425314)

[2. Audit of Equipment 14](#_Toc353425315)

[Introduction to Imaging Facilities 14](#_Toc353425316)

[Imaging and 3D Facilities: A Working Definition 14](#_Toc353425317)

[3D Facilities 15](#_Toc353425318)

[Winchester School of Art 15](#_Toc353425319)

[Archaeology 15](#_Toc353425320)

[Geography 16](#_Toc353425321)

[School of Engineering Sciences 17](#_Toc353425322)

[2D Facilities 18](#_Toc353425323)

[Archaeology 18](#_Toc353425324)

[Findings and Recommendations for Further Work 18](#_Toc353425325)

[3. Data Use Survey 19](#_Toc353425326)

[Methodology 19](#_Toc353425327)

[Results 20](#_Toc353425328)

[Introduction 20](#_Toc353425329)

[3D Data 20](#_Toc353425330)

[3D Data Management: Recommendations 26](#_Toc353425331)

[2D Data 27](#_Toc353425332)

# Introduction

This report contains the findings of a case study which has been undertaken as part of the University of Southampton JISC funded DataPool project. The case study has focussed upon developing improved understandings of data management practice amongst users of raster and 3D data at the University of Southampton. It has also aimed to encourage on-going improvements in data management and to provide recommendations as to how these improvements might be supported at an institutional level.

# Report Structure

The project consisted of three main tasks; each of these will be dealt with separately within this report.

1. Survey of Extant Guidance: As part of the project a survey was conducted of data management guidance available to researchers using raster and 3D data. This part of the report will describe the methodology which was followed in order to identify these resources. It will then go on to describe the resources which were located and will assess the extent to which they represent a comprehensive guide for researchers.
2. Audit of Equipment: A survey of facilities dedicated to the capture and processing of 3D and Raster image data was carried out. The results of this survey were added to existing equipment lists made available through the University of Southampton’s Open Data Service.  This section of the report will describe the methodology used in order to locate these facilities. It will then go onto make recommendations relating to the future development of this resource.
3. Data Use Survey: This section will detail the results of a survey of a sample of users of 3D and raster data at the University of Southampton.  Participants were invited to complete an online survey and were also interviewed. This part of the study sought to better understand how participants respond to the challenges of data sample of users of 3D and 2D data at the University of Southampton.

# Aims, Objectives and Outputs

**Aims:**

1. To better understand how research communities at the University of Southampton are dealing with the challenges of managing raster and 3D data.
2. To use these improved understandings to support improvements in data management
3. To improve communication between researchers in order to encourage research collaboration and resource sharing.

**Objectives:**

1. To conduct a data use survey of raster and 3D data users at the University of Southampton. This will focus on existing approaches to data management and will be based upon interviews and questionnaire data.
2. To produce a comprehensive list of existing guidance documents which will help researchers who wish to develop or to improve their approach to research data management.
3. To audit equipment used for raster and 3D data capture and manipulation at the University of Southampton and to update and increase the coverage of equipment and facilities lists originally created as part of the Southampton-led EPSRC funded Research Facilities and Equipment Sharing Project.

**Outputs:**

1. A report detailing the findings of the data use survey.
2. A web resource which will provide a centralised point of access for University of Southampton researchers to locate data management guidance.
3. An updated list of equipment and facilities related to the capture and manipulation of raster and 3D available at the University of Southampton.

# Survey of Extant Guidance

A key component of this project has been the assembly of existing documents and guides which may assist researchers in developing or refining their approach to RDM. The documents mentioned here relate specifically to the management of raster and 3D data, the list does not include general guidance documentation such as the UK Data Archive’s guide to data management [[1]](#footnote-1)or general advice offered by the Digital Curation Centre[[2]](#footnote-2). The following section will summarise the scope and range of guidance documents available and will then go on to detail each of the documents that has been located.

Methodology   
Guidance documents were located using a range of strategies. The initial list was developed with the assistance of specialists in research data management at the University of Southampton. The list was then augmented based upon discussions with the sample of researchers who participated in the data use survey, the results of which are outlined below. In order to ensure that the list was as comprehensive as possible an extensive web search was conducted to locate any sources of guidance which had not been located using the above methods. Where guidance was unavailable relating to specific disciplines or data types researchers in relevant areas were contacted in order to ascertain whether guidance documents existed.

Introduction to Guidance Documents  
Digital data capture and production has increased rapidly over recent years. This increase in the number of devices and the range of available technologies has led to the need to guide and help users in their daily tasks.   
One of the main institutions in the UK which is dedicated to helping higher education and research institutions in their digital innovations is JISC. They provide wide variety of guidelines relating to the discovery, creation, management and dissemination of data. Another institution providing guidance for managing research data is the Archaeology Data Service (ADS). Their main audience is archaeological in nature but the guidelines are sufficiently general that they can be used in a variety of use cases. As ADS’ main function is to preserve and disseminate the archaeological research data many of their guidelines are produced to raise the quality of the data people/institutions deposit with them.

There are specific resources which deal with metadata. A topic which is probably the most disliked (mainly because of complexity and time requirement) by researchers and therefore very often ignored. There are several standards available depending on data types but often the complexity of the schemas are too complicated to be used without the aid of automated tools. One of findings of the Institutional Data Management Blueprint (IDMB) project was that working with metadata has to be as straightforward process as possible in order  to be used by researchers - IDMB Final Report ([Takeda 2010](#_ENREF_3)).

There are hundreds of guidelines available regarding the best practice in data management for imaging related datasets and workflows. Some of them are highly technical (for example *Guidelines for Handling Image Metadata[[3]](#footnote-3)* by Metadata Working Group), some of them very general (for example *Managing digital images as records[[4]](#footnote-4)* by The State Records Authority of New South Wales, Australia). There are guidelines available for almost any topic related to imaging, the question is, how usable it will be in any particular case.

Workflows surrounding the generation and use of 3D data are changing very rapidly and consequently extant guidance is not always up to date or useful. This is problematic for technical reasons. 3D data have traditionally been generated using specialist equipment such as laser scanners. Increasingly though image data are used to create 3D data. This creates a more blurred distinction between RDM guidance relating to 3D and raster image data, e.g. RTI data management, photogrammetry data management, modern laser scan data management (with associated raster image data).

It is also problematic in terms of engagement. As the use of 3D data enters the mainstream it can no longer be assumed that users have a grasp of technical jargon relating to 3D technologies. Guidance often assumes a higher level of understanding than users have.

Another issue is the proliferation of uses of 3D data. Lots of guidance assumes that certain methodologies are in place. This may no longer be true. For example, the ADS guide to ‘Virtual Reality’ a term which very few contemporary archaeological researchers would use to describe their work.

## Managing 3D Data

The use of 3D data in research is characterised by the diversity of devices, software, formats and workflows which are in use. 3D data frequently forms a discrete component of a larger methodology. A consequence of this heterogeneity is that available guidance tends to be highly specific, referring to specific disciplines, subjects or technologies. Standards, where they are specified, tend to cover specific data use cases (e.g. CAD data) within specific disciplines (e.g. archaeology) although it should be noted that there are very few instances exist of standardisation even at this level. There are very few, if any, sources of documentation which deal with the management of 3D data as a generic category. This situation is in contrast to the treatment of raster data for which general guides to good practice are common, as illustrated in the *Managing Raster Data* section below. This disparity of treatment is an inevitable consequence of the variety of uses to which 3D data are put. The absence of standardised approaches to 3D data management is reflected in the responses given by participants in the data use survey outlined below in section 3 of this report.

It will be noted that Cultural Heritage and Archaeology seem to have produced a disproportionate number of the guidance documents relating to the 3D data when compared to other disciplines. This is certainly due in part to the fact that archaeology and cultural heritage have an inherent focus on material objects. Other factors which may account for this anomaly remain unclear and deserve further study.

In order to ensure the usefulness of this list to researchers descriptions of data type covered and subject area have been included, along with a brief description of the content of the guide.  

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Title** | **Author** | **Year** | **Institutions** | **Discipline/Subject Area** | **Data Types** | **Comments** |
| 3D Laser Scanning for Heritage | David Barber, John Mills | 2009 | English Heritage, Heritage3D | Cultural Heritage | Lidar, Laser Scanning | An Introduction to the technology featuring guidance on project management and data management |
| Metric Survey Specifications for Cultural Heritage | David Andrews (ed) | 2011 | English Heritage | Cultural Heritage | CAD, Laser Scanning | A Guide to 3D documentation of cultural heritage objects with an emphasis on architecture |
| A Guide to LIDAR Data Acquisition and Processing for the Forests of the Pacific Northwest | Demetrios Gatziolis, Hans-Erik Andersen | 2008 | United States Department of Agriculture | Geography | Lidar | A guide to the use of terrestrial laser scanning and the management of outputs as research data |
| Curating Architectural 3D CAD Models | MacKenzie Smith | 2008 | MIT Libraries | Architecture | CAD | A guide to the curation of 3D architectural CAD models |
| Tools and Expertise for 3D Collection Formation | Reinhard Klein, Christopher Schwartz | 2012 | 3D-COFORM, University of Bonn | Cultural Heritage | General 3D Data | A summary of technologies for the documentation of cultural heritage objects |
| The NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials | National Initiative for a Networked Cultural Heritage | 2002 | Humanities Advanced Technology and Information Institute, University of Glasgow, National Initiative for a Networked Cultural Heritage. | Cultural Heritage | General | A guide to good practice for the documentation of all cultural heritage materials including but not limited to 3D data |
| Laser Scanning for Archaeology: A Guide to Good Practice | Angie Payne |  | Archaeology Data Service | Archaeology | Laser Scanning | A Guide to the use of laser scan data encompassing the entire data lifecycle from creation to archiving |
| GIS: Guide to Good Practice | Mark Gillings , Alicia Wise (eds) | 2011 | Archaeology Data Service | Archaeology | GIS | A Guide to the use of GIS data encompassing the entire data lifecycle from creation to archiving. It should be noted that while GIS is not inherently 3D, a significant proportion of the data have a 3D component. |
| CAD: A Guide to Good Practice | Harrison Eiteljorg II, Kate Fernie, Jeremy Huggett and Damian Robinson (eds) | 2011 | Archaeology Data Service | Archaeology | CAD | A Guide to the use of CAD data encompassing the entire data lifecycle from creation to archiving. |
| Creating and Using Virtual Reality: a Guide for the Arts and Humanities | Kate Fernie and Julian D. Richards | 2002 | Arts and Humanities Data Service | Arts and Humanities | General 3D Data with an emphasis on modelled content | A guide to the creation and use of Virtual Reality within a research context, including content on data management |
| Model-Based Engineering Project | Joshua Lubell | 2011 | National Institute for Standards and Technology | Engineering | CAD, General 3d | Details of a project to develop a draft ISO standard for the use of 3D models encompassing Product Manufacturing Information. Although not yet complete the project represents an attempt to impose standards on the use of 3D data in engineering |

Table 1.1 Illustrating major sources of data management guidance for to 3D data

## Managing raster Data

As already mentioned above there is guidance available for almost any task related with managing raster data. There are three bigger topics discussed:

* managing metadata
* image types and compression
* tools for image manipulation and archiving/disseminations

Those three topics can be divided up into data lifecycle as JISC has done on their website:

* find
* create
* manage
* digitisation
* delivery/use

Metadata is big player during the whole lifecycle of the raster dataset, therefore most of the guidelines talks about it. Still, from the experience and from the DataPool interviews it can be seen that often the only metadata people use is the project/object name in file or folder name. Most of the guidelines approach the topic from the collection point of view, assuming that user is using or will set up the system (content management system / digital asset management system / desktop client ) for working with their assets. This kind of guidance is more useful for institutional deciders that lone researchers who would like to maintain their research data.

The best guidelines are for very specific tasks (for example ADS Archiving Close-range Photogrammetric Data[[5]](#footnote-5)), step by step walk through where user does not need to make decisions. Other informal option is email lists.  Recent experience with Britarch mailing list was very good where one member put together very easy-to-follow step by step guidelines about solving particular image archiving related task.

The ultimate end of the line are the guidelines for developers who are working/developing content management systems (see *International Image Interoperability Framework[[6]](#footnote-6)* by University of Stanford and *Guidelines for Handling Image Metadata[[7]](#footnote-7)*, by Metadata Working Group). Usually specification of that kind are very detailed and well written.

To conclude the topic: there are good high quality documentation available about the raster image data management and good guidelines for institutions, but single researcher has to filter out information from the more general guidelines and consult with mailing lists and forums.  
Below some raster imaging related guidance resources are provided with a small comment who could be the target group.

|  |  |  |
| --- | --- | --- |
| **Organisation** | **Document title** | **Comment** |
| JISC | Choosing a System for Managing your Image Collection | Good overview but too technical for beginner. |
| JISC | Systems for Managing Digital Media Collections | Good basics and more info for going deeper |
| JISC | Metadata and Digital Images | Good basics and more info for going deeper |
| ADS | Raster Images: A Guide to Good Practice | Relatively easy to follow |
| ADS | Archiving Close-range Photogrammetric Data | Straightforward guidelines for preparing data for deposition |
| University of Stanford | International Image Interoperability Framework | Very technical, for developers and deciders with technical background |
| GETTY | Imaging related publications | Good selection of online books about imaging. Very accessible |
| National Archives | Digital Preservation Guidance Note 5: Image Compression | Not very technical, useful for all data creators |
| National Archives | Digitised Image Specification | Straightforward guidelines for specific issue. Usable for anyone |
| Metadata Working Group | Guidelines for Handling Image Metadata | For decision makers with good technical background and for developers |
| The Library of Congress | XML schema for encoding technical data elements required to manage digital image collections | For developers |
| The Library of Congress | Sustainability of Digital Formats. Still Images. | For decision makers with good technical background |
| JISC | Britarch mailing list. Often good discussion and relevant links to imaging related guidelines | Informal mailing list. Quality varies. Recently very good “how to” guidelines circulated for researcher how to manage their digital images |
| The State Records Authority of New South Wales, Australia | Managing digital images as records | Very general, for non-technical decider |
| Dublin Core® Metadata Initiative | Dublin Core Metadata Standard | Specification of the standard. Very technical, but useful also for non-techies. |
| IPTC | IPTC Photo Metadata Standards | Specification of the standard. Very technical, but useful also for non-techies. |

Table 1.2 Illustrating major sources of data management guidance for to raster data. (Table will be reformatted in final report to be URLs included)

Findings and Recommendations for Further Work.   
The survey of available guidance indicated that documents were often either targeted at focussed groups of researchers working in a highly specific area or they were highly conceptual and offered more general advice. In both cases the language used was often highly specialised.  The survey found that insufficient introductory content was available to researchers coming to issues of research data management for the first time. Both of these factors have the potential to exclude large numbers of researchers from engaging with issues of data management. The results of the data use survey (detailed in section 3 below) seem to confirm this observation.

The provision of more content which is moderately specific (e.g. at a disciplinary level) and which uses accessible language might enable researchers to locate, comprehend and identify with guidance documents. Successful examples of documents of this type include the Archaeology Data Service’s guides to good practice which are written in an accessible way and which are easily identified as being of relevance by researchers in this discipline. These documents often act as a bridge to more esoteric content, one example being the Archaeology Data Service’s jargon free introductions to Dublin Core metadata standards which are embedded within each of their *guides to good practice[[8]](#footnote-8)*.  

The survey of guidance found that availability of guidance documents was highly uneven with entire subject areas having no guidance aimed specifically at the use of raster or 3D data in that field. Some of the fields engaged in 3D capture on a significant scale have no recognised disciplinary guidelines. Whilst in some cases it may not be practical to produce guidance documents of this type there can be little doubt that in the majority of cases guidance of this type would be extremely useful, particularly as a means of encouraging participation in formalised research data management.

# 2. Audit of Equipment

## Introduction to Imaging Facilities

As part of the Imaging case study an audit of equipment facilities dedicated to the production of 3D and raster data was undertaken. This part of the study was conducted in order to augment existing lists of equipment and facilities compiled by the Southampton-led EPSRC funded Research Facilities and Equipment Sharing Project. The list is available at <http://data.southampton.ac.uk/facilities.html>.

Improved knowledge of available resources and facilities has the capacity to encourage collaboration between faculties and researchers as well as allowing decisions relating to investment in resources to be better informed. Developments in these areas have the capacity to improve the quality of research and will also lead to significant efficiency savings.   
The list as it existed prior to this project has been reliant upon the voluntary submission of data by users; consequently no guarantee of completeness could be made. Investment in equipment and facilities is frequently undertaken at a ‘local’ level using funding from specific projects or departments and has not been added to the list.

Imaging and 3D Facilities: A Working Definition  
In order to conduct an audit of this type it was necessary to define which facilities or devices would fall within the remit of this study. The criteria by which devices were included or excluded from the original list are not explicitly stated on the Research Facilities and Equipment Sharing Project website[[9]](#footnote-9).

Devices dedicated to the capture of raster and 3D data have become increasingly prevalent within academic research at all levels. Consequently it would not have been possible or useful to locate all imaging/3D data capture devices within the University. It was decided that widely available imaging devices such as compact cameras, laptops and mobile devices which incorporate cameras or GPS units would not be included in this study.

The goals of the original study relate to increasing the sharing of technology within the institution and to greater efficiency in investment. In this spirit this audit restricted itself to devices which have a substantial financial value or which are unique or unusual within the University.   
  
Methodology   
A number of techniques were employed in order to locate equipment and facilities within the University. Where possible, staff responsible for equipment were contacted in order to inform them of the study and to acquire addition information relating to the equipment and its availability to other researchers. The following search techniques were employed:

* Use of ePrints (The University of Southampton’s Institutional research repository) to search for publications related to the use of raster or 3D data.
* Conducting broader web searches of University of Southampton research which had a 3D or raster based component.
* Snowball sampling. Researchers within the University who are known to be involved in the capture or use of 3D or raster image data were asked which facilities they were aware of elsewhere in the University.

This combination of techniques was chosen in order to achieve the highest possible degree of coverage. It should be noted however that due to the nature of the study the additions made to the list cannot be guaranteed to be 100% comprehensive.

## 3D Facilities

The following facilities and devices were identified which are used primarily for the creation or processing of 3D data. The devices are listed here by department or school.

Winchester School of Art

* 2 x Makerbot Replicator 3D printer

A Portable 3D printer with a build size of approximately 14cm by 13cm.

Contact: Adam Procter: [adam.procter@soton.ac.uk](mailto:adam.procter@soton.ac.uk)

* Z Corporation ZScanner Handheld laser scanner

A handheld laser scanner for the capture of high resolution 3D data

Contact: Chris Carter: [C.Carter@soton.ac.uk](mailto:C.Carter@soton.ac.uk)

Archaeology

* Konica Minolta Vivid 910 Laser Scanner

Triangulation laser scanner for the capture of high resolution 3D data

Contact: Graeme Earl: [graeme.earl@soton.ac.uk](mailto:graeme.earl@soton.ac.uk)

* Leica GPS1200 Global Positioning System

GPS systems intended for topographic survey

Contact: Tim Sly: [tim.sly@soton.ac.uk](mailto:tim.sly@soton.ac.uk)

* Leica Viva series  Global Positioning System (base station and rover)

GPS systems intended for topographic survey

Contact: Tim Sly: [tim.sly@soton.ac.uk](mailto:tim.sly@soton.ac.uk)

* 2 x Leica SR530  Global Positioning System

GPS systems intended for topographic survey

Contact: Tim Sly: [tim.sly@soton.ac.uk](mailto:tim.sly@soton.ac.uk)

* Leica TC600 Total Station

Reflectorless total station

Contact: Tim Sly: [tim.sly@soton.ac.uk](mailto:tim.sly@soton.ac.uk)

* Leica TCR307 Total Station

Reflectorless total station

Contact: Tim Sly: [tim.sly@soton.ac.uk](mailto:tim.sly@soton.ac.uk)

* Leica TCR703 XR Total Station

Reflectorless total station

Contact: Tim Sly: [tim.sly@soton.ac.uk](mailto:tim.sly@soton.ac.uk)

* Leica TCR805 Power Total Station

Reflectorless total station

Contact: Tim Sly: [tim.sly@soton.ac.uk](mailto:tim.sly@soton.ac.uk)

Geography

* Leica Scanstation C10

Terrestrial laser scanner for capture of large scale 3D data

Contact: Liam Riddy: [L.D.Riddy@soton.ac.uk](mailto:L.D.Riddy@soton.ac.uk)

* Leica Scanstation 2

Terrestrial laser scanner for capture of large scale 3D data

Contact: Liam Riddy: [L.D.Riddy@soton.ac.uk](mailto:L.D.Riddy@soton.ac.uk)

* Leica Flexline TCRP 1205 Total station

Reflectorless Total Station

Contact: Liam Riddy: [L.D.Riddy@soton.ac.uk](mailto:L.D.Riddy@soton.ac.uk)

* Leica GS09 Differential Global Positioning System (x2)

GPS systems intended for topographic survey

Contact: Liam Riddy: [L.D.Riddy@soton.ac.uk](mailto:L.D.Riddy@soton.ac.uk)

### School of Engineering Sciences

* Dimension 1200es 3D Printer

Build size: 254 x 254 x 305 mm (10 x 10 x 12 inches)

Contact: Jim Scanlan: [J.P.Scanlan@soton.ac.uk](mailto:J.P.Scanlan@soton.ac.uk)

* EnvisionTec / Ultra 3D Printer

Suitable for producing useable parts and visualisation, Uses DLP (digital light projection) to cure a choice of high quality materials that are readily available such as ABS-like, polypropylene-like and wax-like plastics

Contact: Kevin Smith: [kevin.smith@soton.ac.uk](mailto:kevin.smith@soton.ac.uk), Mike Street: [mds@soton.ac.uk](mailto:mds@soton.ac.uk), Richard Dooler: [R.J.Dooler@soton.ac.uk](mailto:R.J.Dooler@soton.ac.uk)

* 3D Systems / ZPrinter 650

Suitable for producing parts for visualisation purposes, Uses powder and binders, can print in colour, parts are infiltrated afterwards to make them handleable.

Contact: Kevin Smith: [kevin.smith@soton.ac.uk](mailto:kevin.smith@soton.ac.uk), Mike Street: [mds@soton.ac.uk](mailto:mds@soton.ac.uk), Richard Dooler: [R.J.Dooler@soton.ac.uk](mailto:R.J.Dooler@soton.ac.uk)

* BFB-3000 Plus 3D Printer (x2)

Desktop 3D Printer

Contact: Shoufeng Yang: [S.Yang@soton.ac.uk](mailto:S.Yang@soton.ac.uk)

* ZBuilder/Ultra

Rapid Prototyping Machine

Contact: Shoufeng Yang: [S.Yang@soton.ac.uk](mailto:S.Yang@soton.ac.uk)

## 2D Facilities

### Archaeology

* Reflectance Transformation Imaging kit (x2)

There are two kits for RTI including:

* DSLR camera
* two tripods
* flash
* wireless connectors and remote controls
* other small accessories

Contact: Graeme Earl: [graeme.earl@soton.ac.uk](mailto:graeme.earl@soton.ac.uk)

* GigaPan kit

GigaPan is automated tripod head for capturing panoramas / very high resolution images in controlled way. Each kit includes one GigaPan EPIC robotic camera mount.

Contact: Graeme Earl: [graeme.earl@soton.ac.uk](mailto:graeme.earl@soton.ac.uk)

## Findings and Recommendations for Further Work

In conducting this stage of the research we made several observations which may be considered in the further development of an open list of facilities and devices at the University. These recommendations are not limited to the cataloguing of 3D and raster capture devices but are more generally applicable.

The first and most profound of these was that in order to be sustainable the list should be editable by researchers. Adding data to the list is only currently possible via an email address. A ‘self-service’ system would allow the list to be more easily kept up to date with users able to update and edit the list when necessary. A self-service system would also regularise the collection of data, ensuring that specific data are submitted. Entries in the current list are inconsistent which impedes the usefulness of the list as a means of locating and comparing devices.

Another issue with the list is the nature of the equipment which it contains. Criteria for inclusion are ambiguous and consequently the list is unreliable when assumed to be a comprehensive source. As noted in the introduction to this section it may be necessary to exclude certain categories of equipment which are increasingly prevalent in order to keep the list at a manageable scale. It may not for example be of useful to list every laptop with a camera despite the fact that this can be used as a 3D or raster capture device. It is imperative though that these criteria are made explicit so that the precise nature of the list is clear to users. This will become increasingly important if users are given permissions to edit the list themselves.

# 3. Data Use Survey

The data use survey was conducted in order to improve understandings of raster and 3D data use amongst researchers at the University of Southampton. A particular emphasis was placed upon the strategies employed by researchers in order to cope with the challenges posed by data management.

Participants were encouraged to speak freely about the way in which they managed data and were encouraged to discuss areas which they found to be particularly problematic.

## Methodology

Researchers were selected from the disciplines outlined below using a snowball sampling strategy. Snowball sampling involves contacting an individual or small core of participants involved in the community which you wish to study and allowing them to nominate other participants for the research. These participants can then in turn nominate others until the sample is sufficiently large([Babbie 2010:208](#_ENREF_1)). This technique is acknowledged as being particularly appropriate for the study of populations which may not be easy to identify or access by other means. Utilising the pre-existing contacts of participants meant that a highly interdisciplinary sample, all of whom were conducting appropriate research, could be quickly identified. The inherent bias in samples relating from the use of this technique was not deemed to be unduly problematic due to the limited size of the potential population.

Twenty researchers were identified. Ten researchers were chosen due to the prominent role of 3D data within their research and ten were selected due to their equivalent use of raster data. Participants were asked to fill in a web based questionnaire (see appendix 2) and were then invited to attend a face to face interview during which their answers to the questionnaire were discussed. In the event that they had been unable to fill in the questionnaire the questions were asked for the first time at interview.

As interviews progressed it quickly became clear that the rigid structure of the questions was not viable. The role which data played within the research practice of the participant varied significantly and so it was frequently necessary to deviate from the written questions and follow a more conversational approach.

Interviews were recorded and transcribed and are attached to this report, see appendix x (to be added). The results of the interviews are described below. Participants have been anonymised but certain research groups have been named where they are used as specific case studies. In these cases permission has been sought from those mentioned.

It became apparent at an early stage that there was a high incidence of raster image use amongst the participants that had been selected due to their use of 3D data. The users of Image data were also familiar with, and would often mention 3D data. As a result there were many points made which spanned both groups. These observations will be made first before the analysis continues to look at those matters which are of significance primarily to one group or the other.

## Results

### Introduction

### 3D Data

In addition to the observations made above many areas of interest were identified which relate specifically to the management of 3D data.

#### Infrastructure and the 3D Data Boom

One factor highlighted by the majority of participants involved in the use of 3D data was a rapid expansion in the production of data. To some degree this growth seems to have been facilitated by the recent acquisition of dedicated 3D capture devices. However, in several instances the majority of 3D data was being produced using photogrammetric techniques. This has potential implications for the planning of infrastructure as the capacity to produce 3D data is no longer dependent upon specialist equipment or even upon specialist software. One researcher who was interviewed primarily due to his use of raster imagery had recently begun producing 3D data using a tablet computer and 123DCatch (a freely available photogrammetry application produced by AutoDesk). The use of comparatively cheap devices and free software to produce 3D data widens access to 3D data production enormously. While no figures have been collected relating to the volume of 3D data produced within the University it is clear that this change has the potential to instigate a vast increase in the volume of 3D data produced.

The growth of 3D data, which it is safe to assume is already underway, has serious potential implications for planned infrastructure developments both in terms of the additional storage space which may be required but also in terms of the development and implementation of systems to allow this data to be appropriately curated and archived.

The potential implications of this change are highlighted in the experiences of two research groups at the University.

Research group A has a long history of producing 3D data. Initially this data took the form of 3D models but recently the group’s activities have been increasingly focussed upon the acquisition and processing of high resolution 3D data sets. During this time the curation and archiving practice of the group has changed very little. The projects which the group were involved with were comparatively long, some having lasted more than 10 years but many lasting at least 2 years. The process of data curation used by the group consisted of using hierarchical folder structures stored on the University J drive. Folder structures were designed on a project by project basis meaning that archives were not easily accessible except by those who had created them. For some time now the group have been close to the storage limits of the J drive and have been required to delete large quantities of data on an ad hoc basis in order to make room for new projects. Projects are regularly submitted to a national disciplinary archive. The size of these projects (frequently in excess of 100GB) means that this option is not always available due to cost factors.

This informally implemented policy has several issues associated with it. Firstly, the lack of data standards associated with 3D file formats means that data are often stored in proprietary data formats which is not a sustainable solution. Furthermore, storage in this way limits the extent to which the findability of data with few options available for associating metadata with stored data and no options for searching this metadata.

Participants from this group acknowledged the need for a formalised group-wide research data management strategy which included the destruction of specified content when projects are archived. Researchers also indicated that they would like to use a content management system for 3D data so that metadata could be made searchable and could be associated with files and projects but were reticent about adopting a system which may become redundant in the near future. This uncertainty reflects a broader uncertainty relating to the management of 3D data and the fact that no non-proprietary tools have gained widespread popular usage in the same way that tools such as Flickr, Picasa, Apple iPhoto, Apple Aperture or Adobe Bridge have revolutionised the management and curation of photographic collections.

The group cited time constraints as being the main factor limiting these changes and acknowledged the need to move away from de facto proprietary standards and data management solutions towards sustainable formats and tools. Concern was voiced by researchers from the group that the current system impeded the usefulness of data and that greater accessibility might lead to more regular re-use of 3D content at a local, institutional and global level. Researchers voiced frustration at the costs associated with the procurement of additional storage from the University citing a figure of £1000 per terabyte quoted from the University of Southampton’s Research Data Management website[[10]](#footnote-10). When asked whether they had considered local storage as an alternative to the use of central University repositories users expressed concern about the security of such an arrangement. They demonstrated an acute sense of the financial value of research data and the costs which would be associated with re-recording projects due to drive failure.

Group B faced similar challenges to those of Group A but had responded in a different way. The research group operated a dedicated data capture facility and felt from an early stage that centralised University storage had not represented a viable option. Initially this had been due to limitations in the speed of accessing data via the University network but were described as more recently having had to do with associated costs. The group stated that projects they were involved with frequently generated around 2 terabytes of data with an associated cost of £2000. The group conducted a lot of work for external clients and were concerned that storage in this way would account for 50% of the cost of a project and consequently was not financially viable.   
The solution developed by the group was to invest in a system of local storage. Outputs from capture devices are automatically stored on a bank of hard drives along with metadata which is automatically developed by the device itself. These data are kept in a working directory for a set amount of time before they are automatically archived to the locally managed file store. The system allowed data to be searchable and easily retrieved whenever it was needed. The capture of metadata from the device allowed capture activities to be closely monitored and also allowed accurate calculations to be made regarding the servicing of equipment and the replacement of parts leading to significant efficiency savings.

The group were aware that this system did not offer the level of data security which would be available were they to make use of central University storage but following a cost-benefit analysis felt that this was a reasonable concession. The local store is not treated as a permanent archive and stored data are not guaranteed to clients or researchers using the facility. Users assume responsibility for the security of data. Participants from the group did mention numerous occasions however of data lost by researchers or clients being replaced from the local archive which they maintained.

**Conclusions**

The volume of 3D data which researchers are producing appears to be growing rapidly. All researchers questioned voiced concern that existing approaches to data management might not be sustainable as archives continue to expand. The majority of researchers questioned had little or no plan in place to deal with this expansion.

In part, the problems faced by researchers relate to the limited storage space which they currently have available, however this is only part of the picture. Other problems commonly cited by researchers are outlined below.

#### Data Management: Planning and Standards

Very few of the researchers questioned had written a formalised data management plan for their research. Awareness of guidance documentation was patchy and the majority of participants described research data management in terms of tasks which they encountered within their daily working practice. The upshot of this trend was that researchers were more likely to have formalised approaches to the completion of tasks which related to data capture, data use and data curation than they were for tasks which related to archiving or dissemination. This imbalance was frequently perceived to be a consequence of time limitations. Several participants mentioned that they had very little time available for tasks which did not relate specifically to research and the needs of collaborators and partners.

Where standards were understood they were frequently only partially implemented and only some of the time. A researcher who made extensive use of geographical 3D data cited the example of Dublin Core metadata standards mentioning that while they were committed to instructing their students in the use of metadata they found little time to implement these standards themselves beyond the most basic level.

This situation may be due in part to the limited range of tools available which are dedicated to management of 3D data. While these tools exist, they are frequently designed with specific workflows and applications in mind. Academic research which incorporates 3D data frequently produces highly complex and interlinked data sets made up of multiple 3D and non 3D formats. Furthermore these complex data sets are highly variable, limiting the usefulness of standardised applications or processes. The researchers that had the most success implementing formalised data management strategies were those who produced large numbers of similarly structured data sets, indicating that diversity and complexity in the range of data produced may be a key factor limiting the development of formalised data management strategies. The lack of standards relating to 3D data formats may also exacerbate this situation.

It is also worth mentioning that researchers were far more likely to have created well structured data sets where they were aware that these data (usually as part of a larger project archive) would need to be submitted to a repository for archiving and dissemination. The submission of data to an archive or repository is often built into specific project funding and so researchers are provided with motivation and funding to complete these tasks. Often these data are only structured at the point of submission. This raises the key issue of funding. Researchers were acutely aware of the costs implicit in investing time in the development and implementation of data management strategies. Where these funds were made available through funding from grant awarding bodies researchers were likely to spend more time concentrating on these issues.

**Conclusions**

Many researchers had a very limited awareness of how data management tools might help them to structure and organise their data. Clearly, this situation might be remedied through the provision of additional training; the case for this is described in the *guidance* section immediately below. However, it should be noted that time limitations were a near universal factor in the decision of researchers not to invest in the development. If formalised data management strategies are to become more common it seems likely that guidance will need to be made more readily available but also that a sustainable model for funding this work will need to be considered.

#### Guidance

Awareness of guidance amongst researchers using 3D data is mixed. Several researchers were aware of guidance issued by bodies such as the Digital Curation Centre and the UK Data Archive but awareness of guidance and training provided by the University was very limited. None of the researchers questioned had utilised guidance provided by the University of Southampton’s Research Data Management guide hosted by the library[[11]](#footnote-11). None were aware of the work of the Software Sustainability Institute in supporting the use of software in academic research[[12]](#footnote-12). Both of these organisations offer training and consultation as well as written materials.

The limited awareness or use of guidance documentation was also the norm for guidance relating specifically to the management of 3D data or to specific disciplines. It is likely that guidance dedicated to the management of 3D data is perceived as being esoteric and too highly specific. As evidenced in Section one of this report, very little advice exists relating generally to the management of 3D data. While many of the points raised in more specific documents may be transferrable their specialised nature ensures that they may be hard to locate and that they may not be seen as a likely source of insight.

Discipline specific guidance is less widely available with only archaeology and cultural heritage offering a wide range of guidance. There was little evidence for the wide use of these resources within the disciplines towards which they were targeted. This may be in part due to the difficulty of keeping these resources up to date but it seems unlikely that this is the only reason for their limited adoption. One researcher noted that these resources were widely used for teaching purposes but that they were very infrequently used by researchers. This may be due in part to the limited time available for data management activities mentioned above.

**Conclusions**  
The heterogeneous and highly dynamic nature of research practice incorporating 3D data has created a situation in which the provision of guidance documents has been limited. However a range of guidance documentation which is aimed at the research community more generally is of great potential value to the 3D research community and is currently under used.   
It seems to be the case that the limited use of these materials has more to do with a lack of awareness than it has to do with a lack of suitable material. It is probable that greater communication of available resources and perhaps even provision of training relating the effective use of these materials would increase their use among the research community. This would be an inexpensive method of improving the standard of data management practice among the 3D data research community.

Without further research taking place it is difficult to know whether the limited time which researchers feel they have available to deal with data management would prohibit the usefulness of these efforts. It seems likely however that improvements could be made to existing practice simply by encouraging researchers to locate and consult existing guidance.

#### Informal Data Management Solutions

Participants who were selected based upon their use of raster data made extensive use of informal systems for managing and curating their data. This trend was far less common amongst the community of researchers using 3D data. Notable examples from the community of raster researchers included the use of cloud based media sharing platforms such as Flickr and Dropbox to curate data ([Beale, Hitchcock et al. 2013](#_ENREF_2)). This trend was far less prevalent amongst users of 3D data. The reasons for this disparity may be complex but are likely to have to do with the limited availability of cloud based media sharing utilities designed with 3D data in mind, the limited uptake of these utilities within the broader population and the complex nature of projects which incorporate 3D data.

Limited uptake of cloud based media sharing platforms amongst users of 3D data has led to the adoption of less agile systems for data curation and sharing. Users referred frequently to the use of the University’s networked J drive as a means of sharing data although it was noted that this did not allow for external sharing to take place. Participants also made use of Dropbox or large file sharing sites such as YouSendIt or Sendspace in order to share files with external collaborators. Where these sharing platforms were used for 3D data they tended not to be utilised for anything other than the most basic curation tasks such as sharing specific files with individual collaborators. This stands in stark contrast to the use of these systems for the curation of raster data. In these instances cloud based media sharing platforms were used at various stages of the data lifecycle in order to curate, disseminate, publish and document research.

**Conclusions**

It is likely that this disparity has to do primarily with the relative lack of cloud based media sharing platforms dedicated to 3D data. While examples of such systems exist they are not nearly so widely used or widely understood at the raster equivalents. The use of these systems by users of raster data tends to have been driven by the widely perceived usefulness of these systems and familiarity born out of extensive use in the personal sphere.   
  
Communication and collaboration   
Issues of collaboration and communication were discussed by all researchers questioned. Their comments and observations fell into three distinct categories which can be loosely categorised as having to do with:

1. The challenges of managing data relating to collaborative research.
2. The need for and potential of collaborative data management planning.
3. Sharing of equipment
4. Collaboration is common amongst the researchers who were questioned and takes many different forms. Most of the researchers we talked to had collaborative relationships with other researchers or groups either within the university or elsewhere.  
   Many researchers noted that collaboration posed unique challenges in terms of data management but some had begun to identify benefits emerging from collaborative data management.

In terms of challenges posed by collaboration the majority of participants highlighted issues relating to infrastructure. The lack of an institutional system for sharing data externally was seen as being problematic and has led to a widespread reliance upon third party file sharing applications such as Dropbox, YouSendIt and SendSpace. Researchers tended to use the J drive to manage internal collaboration. A desire was expressed by many researchers to have a filestore which could be accessed by researchers from external institutions but which could act as an active research filestore. This would be similar to the J drive but would be accessible without the need for a University of Southampton filestore.

Several researchers had found that the process of sharing data had led to improvements in data management practice. One notable example was a collaborative venture between researchers at Winchester School of Art and Archaeology. The project was created in order to promote collaborative production and processing of 3D data based upon the idea that 3D data collected by each group had the potential to be of value to the other. The collaboration has led to the free sharing of 3D capture devices and 3D printing facilities as well as training in the use of these technologies. This equipment collaboration has also led to the collaborative development of data management methodologies which will allow data to accessible, findable and intelligible to researchers working in different fields. Researchers from this project were also keen to note that this equipment and infrastructure collaboration had acted as a catalyst for the development of collaborative analysis, fieldwork, exhibitions and publications

1. Several researchers who were questioned mentioned that it might be useful to have a means of finding out how other people had coped with the challenges of data management. Very few of the participants had made their data management strategy available to others.
2. There was an extremely low level of awareness of specialist equipment held by other researchers or research groups. It seems likely that this limited awareness has had the effect of inhibiting collaboration which may otherwise have occurred. There is also potential that it might lead to inefficient investment in equipment. A lot of the equipment which researchers described was used only at intervals and might have been used by others within the University. Attitudes towards the sharing of equipment were generally very positive especially where the possibility of reciprocal arrangements was discussed. As described above, researchers at Winchester School of Art and Archaeology have described the ‘higher level’ collaborations which emerged as a result of equipment sharing collaborations. Consequently it seems likely that more efficient equipment sharing has the capacity to:

* Improve the efficiency of investments in equipment
* Encourage researchers to invest in a wider variety of equipment (due to greater awareness of duplicating investments made elsewhere)
* Catalyse collaborative research

### 3D Data Management: Recommendations

#### Management and Planning

Channels should be developed which allow researchers to share data management strategies and to communicate with other researchers who have already successfully implemented and shared data management strategies.

This system could be usefully upgraded to incorporate a means of groups feeding back useful data such as the volume and type of data that they are currently holding. Where possible the collection of this data should be automated to ensure that it is accurate and up to date. This would allow more efficient management of existing storage and would also allow for effective planning and investment in storage.

#### Guidance

Available guidance documents are very often inaccessible to researchers who are not acquainted with research data management. Guidance can be impenetrable to researchers unused to specific terminology or technical styles of writing and are often difficult to locate. In order to address these problems guidance documents should be produced which:

1. Provide a basic introduction to the key concepts of sound data management. These should be highly accessible and may achieve greater impact if they are written on a discipline by discipline basis.
2. Allow researchers to locate guidance which is appropriate to their subject are and/or to the data types which they work with.

These documents will be inexpensive to produce and if widely disseminated and publicised will provide researchers with a guide to better data management. Dissemination of these documents could be supported by a series of workshops which introduce the principles of sound data management practice and also make researchers aware of the range of advice and help available to them.

#### Equipment and Investment

The sharing of equipment was central to the development and maintenance of meaningful inter-disciplinary collaborations amongst 3D data users. Several initiatives exist across the University of Southampton which bring together users whose research interests involve the use of 3D imaging. These include a 3D printing group, collaborations between Arts and Humanities groups and research groups within Engineering and also collaborations between Archaeology and Winchester School of Art as outlined above.

Recommendations relating to the development of the equipment list are made in section 2 of this report. These changes would go a long way towards making researchers aware of the range of equipment held by research groups at the University of Southampton. In addition to these efforts it would be very effective to support these collaborations in other ways. Successful examples which might be extended include the development of University Strategic Research Groups[[13]](#footnote-13) and the recent awarding of small grants to PhD students by Southampton Digital Humanities[[14]](#footnote-14).

These efforts might be effectively supported by a series of sandbox events which are open to PhD researchers and staff at the University and which result in the provision of small grants for research projects which make innovative use of equipment.

### 2D Data

Data Use Survey brought up two very distinctive usergroups:

1. units providing professional services (such as Library, former GeoData Institute and Microscopy Centre)
2. lone researchers or small research groups

Following discussions are based on this division.

Data Management  
In group one workflows were very well worked through and established. For example widely accepted standards were used for metadata collection and ingestion. All users in this group were aware of all the standards available for their data types and used them as part of their workflows. Same applies to storage and archiving. All of them use backup systems to make sure that the data is safely stored. In some cases it was provided by iSolutions and in one case the system was using the J-drive. In this case all the data was copied from the imaging device to the lab’s computer which automatically made it available to the users through J-drive. This is a good example where existing system is used to integrate into lab’s data management workflow. It is always easier to adapt something which one is already familiar with. In other case iSolutions was asked to evaluate the system for particular data management workflow. Again, study participant was very positive about the support provided by iSolutions.

Second group of users did not follow any particular workflows. Most common use case was where data is copied from the capture device into user’s hard drive and backed up in external hard drive. Sometimes copy was kept only on external hard drive. In one case all those hard drives were kept at the same office, so in worst case scenario all the data could be under risk of damage or loss. Network drive (J-drive) was used for project which had folder already created but even then not always. In one particular case J-drive had a copy of original files but all the processing was done in local hard drive because of the network speed (some procedures are extremely slow when files are in network drive) or/and compatibility issues with processing software (several software products does not let you work with mapped drives or cause crashes). As one of the interview participant reported this kind of workflow has created issues with versioning, so the archive on the J-drive was not complete.

When it comes to metadata, only one user kept the spreadsheet with the metadata about the image files he had. Other users used only file and folder structure to store meaningful information about their data. There were cases where during the capture process log book was kept what was captured but this information was then converted into folder structure.

Informal Data Management Solutions  
There are many tools available for image data management. Online media sharing services like Flickr or Picasa have become very popular. Also general file sharing services such as Dropbox have gained popularity as platform for sharing research data. Main reason behind that, and this was confirmed by several interview participants, is the usability and accessibility, something which most of the university provided services lack.

When talking about raster data specific workflows, and mainly its delivery or dissemination phase, then tools like Flickr become even more important. It is not just the tool for sharing the files but it becomes a tool to work and analyse the data with.

Infrastructure  
Almost all of us use daily one or more raster data capture devices: mobile phones, computers with webcams, “point and go” or SLR cameras. The greater accessibility of devices has increased also the amount of data produced with them. As it will be discussed later in most cases the data ends up on user’s computer local hard drive. Leaving aside all the data management issues it should be pointed out that computers themselves are part of the imaging facilities as they are used to process and generate new data. In several cases described by interviewees data processing was not possible or was extremely time consuming. The reason was (at least described this way by interviewee) not the bad hardware but the University’s standardised workstation software / network setup. In both cases those workstations were in dedicated computing labs for working with different kind of imaging. Point to make here is if university should support special setups for dedicated labs without extra (licensing) fees for the research group or researcher.

Guidance  
As above we can divide awareness of guidance also into two groups. First group was aware of all the industry and research community standards; other group relied more on colleagues help and advice.

Communication  
Communication seems to be the key issue in data management and institutional support. It came out clearly from the interviews that if one asks for support, one probably gets it. For example in two cases users were aware that the default network drive space provided by iSolutions was too small but they never asked for more. Sometimes it is not very easy to know where to ask though.

Another point made by interviewee was that university should make some feedback/guidance related self-service tools available to let users to some information to be updated by themselves, rather than waiting while the temp-worker will process the service-line email.

#### Conclusions

Many of the key finds of the raster data management are very similar to any kind of other data management workflows. The main difference is the data visualisation (image viewers etc.). First point to make is that University is probably not able to compete with commercial service provider (such as Flickr, Picasa etc.) on that, rather, use of this kind of services should be written into project's, research group’s data management plans. Then informal becomes formal.  
Second point to make is that University should provide more support on understanding the data management and how to prepare the data management plan. Quite often this is seen as something you “have to do” rather that this is something useful.

Third point follows from previous, the need for better communication. Having guidelines online is not enough. It has to be communicated properly to undergraduate level students as well as to PhD students and staff. Also, tools to let users to help other users could be provided by University.

As pointed out in previous paragraphs we can see the clear division in between groups using service oriented approach and individual researchers. When service is provided to the customer the requirements for the outcomes are always higher and therefore the data management is taken much better care of. This could be also something related to communication that researcher should change their attitude towards their work and see it as the work for the customer (taxpayers, funding bodies).

Babbie, E. R. (2010). The practice of social research. Australia ; United Kingdom, Wadsworth Cangage Learning.

Beale, G., S. Hitchcock, et al. (2013). Supporting Data Management for 3D and Raster Data: Lessons learned from the DataPool Project. IS&T Archiving Conference, Washington DC, IS&T.

Takeda, K. (2010). Institutional Data Management Blueprint Project: Initial Findings Report, University of Southampton

JISC.

1. <http://data-archive.ac.uk/create-manage> [↑](#footnote-ref-1)
2. <http://www.dcc.ac.uk/> [↑](#footnote-ref-2)
3. <http://www.metadataworkinggroup.org/pdf/mwg_guidance.pdf> [↑](#footnote-ref-3)
4. <https://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/guidance/guidelines/guideline-25/managing-back-capture-digitisation-projects/11.-managing-digital-images-as-records> [↑](#footnote-ref-4)
5. <http://guides.archaeologydataservice.ac.uk/g2gp/Photogram_4-1> [↑](#footnote-ref-5)
6. <http://lib.stanford.edu/iiif> [↑](#footnote-ref-6)
7. <http://www.metadataworkinggroup.org/pdf/mwg_guidance.pdf> [↑](#footnote-ref-7)
8. <http://guides.archaeologydataservice.ac.uk/> [↑](#footnote-ref-8)
9. <http://www.southampton.ac.uk/research/centres_facilities/facilities_equipment_database.shtml>   
    [↑](#footnote-ref-9)
10. <http://www.southampton.ac.uk/library/research/researchdata/storage_options.html> [↑](#footnote-ref-10)
11. <https://www.southampton.ac.uk/library/research/researchdata/> [↑](#footnote-ref-11)
12. <http://www.software.ac.uk/> [↑](#footnote-ref-12)
13. http://www.southampton.ac.uk/multidisciplinary/usrgs/index.page? [↑](#footnote-ref-13)
14. http://digitalhumanities.soton.ac.uk/small-grants/ [↑](#footnote-ref-14)