

# EXPERIMENTING WITH THE TRIAL OF A RESEARCH DATA AUDIT: SOME PRELIMINARY FINDINGS ABOUT DATA TYPES, ACCESS TO DATA AND FACTORS FOR LONG TERM PRESERVATION

*Panayiota Polydoratou*

University College London, Library Services  
DMS Watson Building, Malet Place, WC1E 6BT, UK  
e-mail: [p.polydoratou@ucl.ac.uk](mailto:p.polydoratou@ucl.ac.uk)

## **Abstract**

Developing systems and services for the effective and efficient management of research data as well as addressing issues around their long term curation is an area of increasing activity in UK Higher Education. This paper discusses some preliminary results from a questionnaire survey, conducted as part of the trial implementation of the Data Audit Framework Methodology at University College London (UCL). Fifty seven (57) academic and research staff from 5 designated departments and an interdisciplinary research centre provided information about the nature of their research and the types of primary research data they produce. The survey explored factors that could impact on access, use and preservation of such data. The preliminary results indicate that researchers recognise the potential usefulness of such data for other researchers as well as their long term value. Retaining primary research data after the end of the funding period and re-using them for initiating further research are practices already acknowledged. However, ownership, copyright and restrictions on access to research data can be hazy areas for academic and research staff and require further investigation, advice and support. The value of primary research data appears to be closely linked to the context within the data which were generated.

**Keywords:** primary research data; data audit; data management; survey.

## 1. Introduction

The outputs of research activity are published records, most commonly in the form of journal articles, books, reports, etc. However, the vast majority of the data that are produced during the research process never reach publication stage. There is an increasing recognition of the potential benefits of such data both to the wider research community and to society in general.

Developing systems and services for the effective and efficient management of research data as well as addressing issues around their long term curation is an area of increasing activity in UK Higher Education. Recent examples of such activity include projects that currently run under the Joint Information Systems Committee (JISC) Digital Repositories programme 2007-2008 such as the Data Audit Framework Development and a further study for identifying the benefits of curating and sharing research data, and the joint RLUK-RUGIT UK Research Data Service Feasibility study. Moreover, studies of the complex relations and issues around the generation, management, curation, use and value of research data (Fry and colleagues, 2008; Swan & Sheridan, 2008; Lyon, 2007; NSB, 2005; MacLeod & Childs, 2003) have already started to pave the way for further research in this area.

The UCL Data Audit Framework pilot implementation project (UCL DAF) is part of the activity in this area.

## 2. UCL DAF project - Background

University College London is a global research university which encompasses 200 years of discovery and achievements and aims to transform the world. Research has been a core activity at UCL since the university's foundation in 1826 and is at the heart of its future strategy<sup>1</sup>. UCL counts amongst its strengths more than 4000 academic and research staff of which – including 677 professors – working at the forefront of their research fields<sup>2</sup>.

During 2006/2007, UCL staff have successfully secured research funds and excelled in scholarship. Income generated by research funds and grants accounted for one third of the overall UCL income. Recognition for the advancement of scholarship includes fellowships in Royal Societies and Academies, professional associations as well as 20 Nobel Laureates up to now that have been awarded to UCL staff since 1901, most recently to Professor Sir Martin Evans in the field of Physiol-

---

<sup>1</sup> UCL vision and strategy. Information available at: <http://www.ucl.ac.uk/research/vision/>

<sup>2</sup> UCL facts and figures. Information available at: <http://www.ucl.ac.uk/research/facts/>

ogy of Medicine, in 2007. Furthermore, recognition at an international level is demonstrated in the Times Higher Education (THE) rankings that list UCL 9<sup>th</sup> of the world top 200 universities<sup>3</sup>. Outputs from a recent bibliometric study placed UCL in the 2<sup>nd</sup> place of most productive research institutions in Europe, 3<sup>rd</sup> place as the most cited university and the UK University most cited by health researchers (van Raan, 2008).

The diversity, volume and breadth of the data produced by UCL scientists are further enhanced by the merits of interdisciplinary research which, UCL strongly promotes. It often involves collaboration with other universities and at international level and is made evident by the number and cross boundaries of the centres and institutes at UCL. Further information can be found at: <http://www.ucl.ac.uk/research/departments/themes>

UCL Library Services is in the core of the network that supports the university's vision. The Library states its commitment to support excellence of teaching, learning, research and clinical practice by providing a high quality, integrated and innovative service<sup>4</sup>. Furthermore, and in compliance with advances in the provision of access to information, UCL Library Services is committed to the responsible, long-term stewardship of all its digital assets, whether born-digital or created through digitisation. The Library works with the Information Strategy Committee and the Information Systems and Media Resources to align the development of portals with existing tools to realise our mission. In further support, an established Working Group on Digital Curation:

- leads on digital curation issues for UCL
- identifies best practice in the curation of scholarly and administrative materials in digital formats, including primary data
- disseminates best practice in digital curation across UCL
- maps the growth of digital objects in UCL and prepare UCL to undertake their long-term digital curation
- sponsors and pilots new initiatives, technology and processes in digital curation within UCL

Participation in the Data Audit Framework Development project and implementation of the methodology falls exactly in line with UCL's mission and commitment to support research and teaching and the recommendations made by Liz Lyon (2007) in the JISC commissioned report "Dealing with Data". By participating in

---

<sup>3</sup> Times Higher Education (2007). World University Rankings: top 200 world universities. Information available at: <http://www.timeshighereducation.co.uk/hybrid.asp?typeCode=144>

<sup>4</sup> UCL Library Services e-strategy. Information available at: <http://www.ucl.ac.uk/Library/e-strategy.shtml>

the DAF project, UCL aimed to gain an understanding of: a) what primary research data are held at designated faculties and departments, b) how the data are managed, and c) where the responsibility lies for their long term curation. Specifically, the project's aims and objectives are described in the following section.

### **3. Aims and objectives**

The UCL Data Audit Framework project was one of the four pilot/exemplar implementations of the Data Audit Framework. The aims of the UCL DAF project were to contribute to the iterative development of the Data Audit Framework and collate information about the research assets generated, held and managed by academic and research staff at UCL. In particular, the objectives were:

- explore the implementation of the methodology developed by the Data Audit Framework Development project for auditing research assets at Higher Education Institutes
- document, discuss and report issues and lessons learned from the pilot implementation with the members of DAFD and the other pilot implementation projects
- collate information about research assets and data management practices at designated departments/centres/institutes at UCL and
- share the findings with the academic community and beyond.

### **4. The Data Audit Framework Methodology**

The development of the Data Audit Framework (DAF) methodology has been led by the Humanities Advanced Technology and Information Institute (HATII) in conjunction with the Digital Curation Centre (DCC) supported with JISC funding. The methodology can be accessed and downloaded from the DAF website (<http://data-audit.eu/methodology.html>). In brief, it recommends that audits of research data proceed in four stages:

- 1) Planning the audit (Appoint auditor, establish business case, research the organisation, set up the audit, etc.)
- 2) Identifying and classifying assets (Collect information about the organisation's holdings, use methods such as surveys and interviews to identify and classify assets, prepare the inventory, set up meetings to assess findings, etc.)
- 3) Assessing management of data assets (prepare and conduct interviews to assess vital assets, etc.) and
- 4) Reporting and recommendations.

## 5. Scope and limitations

The project aimed to trial the DAF – four stages – methodology (planning the audit, identifying and classifying assets, assessing management of data assets and reporting and recommendations) across a range of departments and an interdisciplinary research centre at UCL.

However, after taking into consideration a) the breadth and volume of research data that a research-led higher education institution is likely to generate and b) the time restrictions of a pilot project, it was decided to restrict the audit to primary research data only. For the purposes of the project, a definition and understanding of the characteristics of primary research data across disciplines and practices was as follows:

Primary research data are data produced within the timeframe of a project/research work/lifetime. It is unprocessed (often referred to as raw data), original, generated by machines or humans and is regarded as the core of any research activity.

Also, this paper presents and discusses findings that relate to stage 2 of the methodology: identifying and classifying data assets only. The reasoning and our implementation approach are discussed in the following section.

## 6. Approach on the implementation of the DAF methodology

Overall, the project aimed to follow the 4-stage methodology, with allowances to be made for the short timeframe and its pilot nature - some selectivity in methodology had to be exercised in order to ensure useful outcomes for both the project partners and the UCL academic community.

### 6.1. Target population

UCL counts amongst its strengths more than 4000 academic and research staff - including 677 (approximately 17%) professors – working at the forefront of their research fields. The following table presents the number of UCL academic and research staff by scientific field and provides some indication of the UCL research activity by faculty and academic departments . Further information about research outputs at UCL is documented in the UCL Research Publications database (<http://www.ucl.ac.uk/research/publications/>) and UCL Eprints (<http://eprints.ucl.ac.uk/>).

Research field	Number of Staff <sup>5</sup>	%
Arts & Humanities	228	5.7
Biomedical Sciences	1818	45.4
Built Environment	130	3.3
Engineering Sciences	366	9.1
Laws	62	1.5
Life & Medical Sciences	595	14.9
Mathematical & Physical Sciences	478	11.9
Social & Historical Sciences	324	8.1
Total	4001	100

Table 1: Number of academic and research staff by scientific field.

The audits were planned with the target population of academic and research staff in mind, including those holding honorary and affiliated agreements with UCL. Five departments and an interdisciplinary research centre<sup>6</sup> were selected representing a range of scientific fields and interdisciplinary research. Specifically, the departments and centre that were approached are:

- 1) Department of Scandinavian Studies (Arts & Humanities)
- 2) Institute of Archaeology (Social & Historical Sciences)
- 3) UCL Interaction Centre (Interdepartmental and cross-faculty research)
- 4) Department of Language and Communication (Life & Medical Sciences)
- 5) Department of Speech, Hearing and Phonetic Sciences (Life & Medical Sciences)
- 6) Department of Physics and Astronomy (Mathematical & Physical Sciences).

## 6.2. Timing

The first stage of the DAF methodology required careful planning of the audits, substantial research and close collaboration with key contacts at the prospective audited departments. Timing, clear scope and an institutional requirement for an audit were considered crucial factors for success. Picking the right time for the audits could ensure staff engagement, completing the audits successfully and pro-

<sup>5</sup> UCL facts and figures. Information available at: <http://www.ucl.ac.uk/research/facts/>.

<sup>6</sup> Some information about the research outputs of the audited departments is listed in Appendix B of the forthcoming UCL DAF project final report. The report will be available at the UCL Eprints (<http://eprints.ucl.ac.uk/>).

duce useful recommendations about the management of research assets. However, the launch of the project coincided with the beginning of the academic term at UCL, a very busy time for both academic and research staff.

Taking into consideration the short duration of the project and limitations imposed by the timing of the project’s launch, it was decided that a questionnaire survey could facilitate quick capture of data at this point and perhaps appeal more than face to face interviews.

The questionnaire was designed around the Form 2 (inventory of assets; available at [http://www.data-audit.eu/DAF\\_Methodology.pdf](http://www.data-audit.eu/DAF_Methodology.pdf), page 43; screenshot shown in figure 1 below) of the methodology. It comprised three sections: section 1 (brief characteristics of the participants/nature of research activity), section 2 (data types, characteristics of research, attributes of primary research data), section 3 (brief description of primary research data to match information requested in Form 2 of the methodology).

The online questionnaire survey ran for 3 weeks and offered an incentive of £50 prize draw (Amazon vouchers) to encourage participation. The survey aimed at serving the following purposes: raise awareness about the project at the designated departments, gather information about types of research assets within discipline, facilitate quick completion of Form 2 of the DAF methodology, and act as means to plan future interviews.

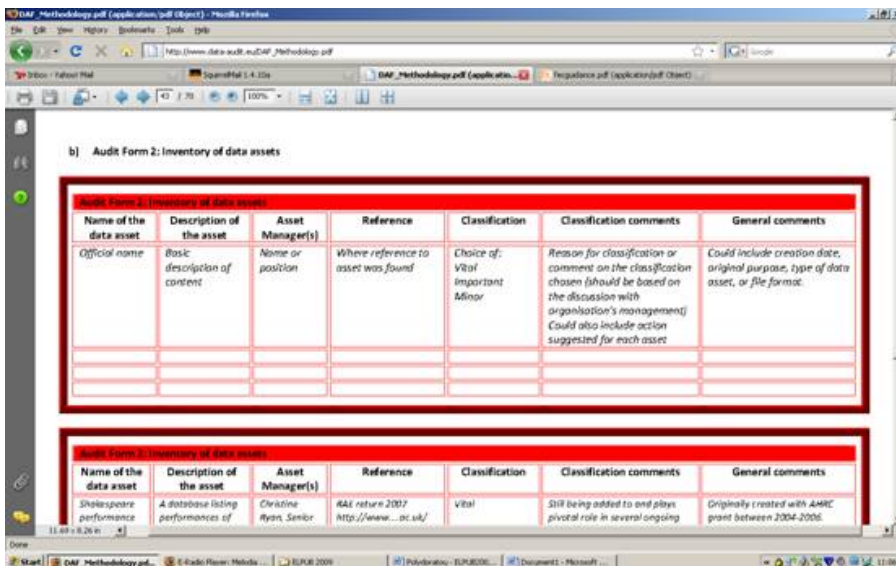


Figure 1: Form 2 - Inventory of Assets.

## 7. Results

One hundred and ninety two (192) people<sup>7</sup> were invited to participate in the questionnaire survey which was administered using the e-inform software. The survey ran for 3 weeks and received responses from 57 people, approximately ~30% of those contacted and collected 32 examples of primary research data. In addition to the survey, thirty (30) people were interviewed (11 volunteered via the survey). The response by audited department is shown in Table 2.

Department/Centre	Number of staff contacted	Survey response	Interviews
Institute of Archaeology	71	25	13
Department of Scandinavian Studies	10	-	6
UCL Interaction Centre	14	5	3
Physics and Astronomy	68	13	4
Research Departments of: Language and Communication Studies and Speech, Hearing and Phonetic Sciences	39	14	4
Total	192	57	30

Table 2: Number of staff who participated in the trial of the DAF.

### 7.1. Nature of research

The participants in the survey were invited to provide some information about their membership in research groups, role at their departments and current involvement in research activity. The majority (42 people) described themselves as academic staff at various grades on their scale (lecturer, senior lecturer, reader, professor, etc.) while response received by an honorary member of staff, 2 Ph.D. students and staff employed in commercial research (for example, commercial archaeology). Almost 90% (51 people) denoted that they were actively involved in research.

The respondents were also invited to describe the nature of their research. In particular, they were offered the option of choosing between computational, experimental and observational research as well as specifying other research. The

---

<sup>7</sup> The Department of Scandinavian Studies decided to opt out from the questionnaire survey and participate in interviews only. Therefore, the number of staff at the department is not included in the questionnaire survey sample mentioned here.



nature of data and their origins could pose particular considerations when it comes to preservation, retention and long term curation. The respondents indicated involvement in all three types of research (response shown in Figure 2) while eight (8) people provided the following further specifications:

- Primarily theoretical (2)
- It is part of all the items above, except for computational
- Interdisciplinary between experimental and observational
- Archaeological survey and excavation
- naturalistic
- Microscopy, chemical analysis; logical thinking; comparison to other / published material publication and analysis
- All three in equal measure.

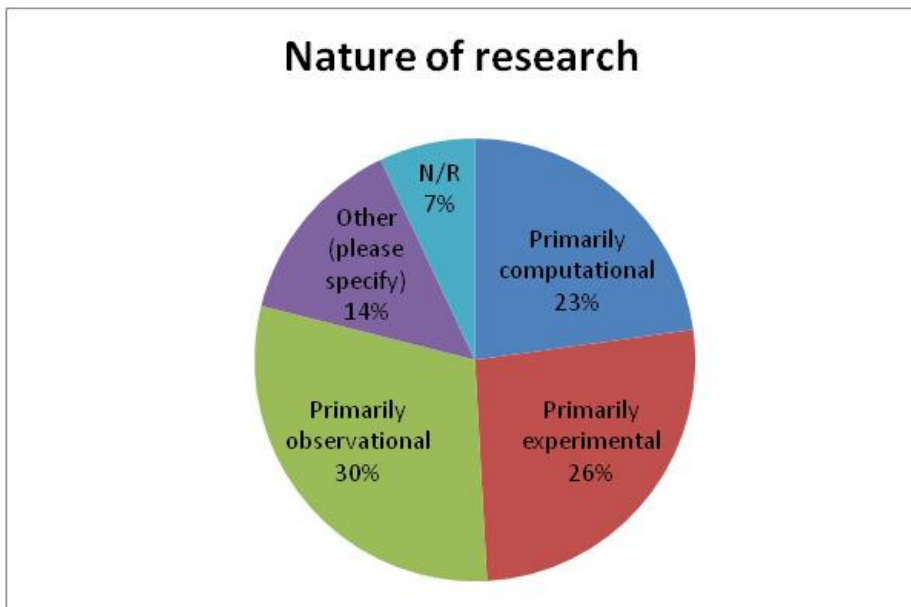


Figure 2: Nature of research.

## 7.2. Type of research data

Primary research data come in various formats and can be quantitative or qualitative. Being aware of the types of research data that are generated could provide some indication about their assorted formats, storage requirements and technology reliance. The members of staff at the audited departments specified that numeric data (35 replies), textual data (24 replies) and databases (20 replies) were the three most likely types of data that their research would generate. Images and au-

dio visual material followed closely. Other types of research data that the respondents specified were: artefacts and text, spatial datasets (GIS layers) and field data (Figure 3).

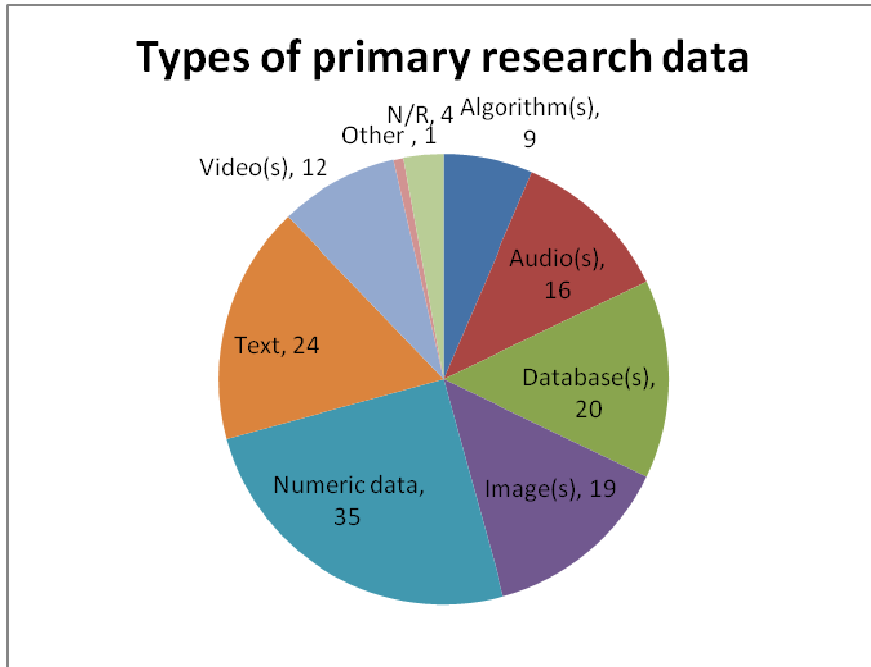


Figure 3: Types of primary research data.

In addition to this question about primary research data types the respondents were asked to provide a brief example of their own research data by filling in the corresponding Form 2 of the DAF methodology (inventory of assets). The information that was requested to accompany those brief examples was: name of the data asset, description of the asset, owner, storage location, classification (in terms of usefulness to the institution – minor, important, vital) and classification comments. The intention was to generate a brief inventory of assets based on user feedback and also to test initial reaction to the proposed classification attributes. Although raw data do not carry an intrinsic value until they can be exploited and turned into information (Buchanan and Gibb (2007)) deciding on criteria and/or prospective indicators of value is a complex matter.

Although the examples provided by the respondents will not be discussed in detail in this paper, perhaps an interesting observation that can be made at this point is that the examples: a) listed at least two different types of data produced by the same researcher, b) referred to collections of data or combinations of data types

rather than single data sets and c) researchers appeared to regard data as material linked together by various associations. Those associations referred to as contextual information and it highlighted the value staff assigned to their data. This observation illustrated in the table below shows examples by researchers from 3 different scientific fields:

	Example 1	Example 2	Example 3
Name of research data asset	motion capture data	<i>[Information omitted]</i> Glass	<i>[Information omitted]</i> project data
Description of the research data asset (brief description of content, purpose, etc.)	numerical description of the movement of a person	images, micrographs, chemical analyses	digital video tapes and word transcripts of conversations – for analysis via repeated viewing, to explore communication strategies used by people who have language difficulties caused by <i>[information omitted]</i>
Owner(s) (e.g. specific researcher, University, funders, etc.)	myself	myself, and co-operating archaeologist	UCL
Location of storage (where are the data stored/kept, e.g. network drive, CD, etc.)	personal laptop and personal desktop	hard drive	external hard drive (mpeg), filing cabinet (digital tapes), network drive (word files of transcripts)
Classification (Minor, Important, Vital)	Vital to my own research	useless to UCL; vital to me if I want to publish	important
Classification comments		I was unsure about quite a few questions what they meant [...]	a unique archive – this kind of data is time consuming to collect and is of great potential use to other researchers

Table 3: Examples of primary research data.

### 7.3. Characteristics of research activity

The participants in the survey were invited to select from a list of options those that best applied to their research activity. The list aimed at gathering information about various factors that could potentially impact on access, use and preservation of primary research data; for example, requirements imposed by the funders of the research (perhaps regarding retention, access to data and/or preservation of research outputs), or specific agreements between the different collaborators in the case of interdisciplinary and across institutions research (e.g. copyright, ownership of data, etc.) as well as those imposed by the nature of the research itself (e.g. dealing with sensitive data).

More than half of the respondents (32) indicated that their research is most of the time externally funded and involves collaborations with other researchers/institutions outside UCL. Working or having to deal with sensitive data did not seem to apply to almost half of the respondents. Furthermore, about two thirds of the replies indicate that respondents rarely contribute to third party databanks or they felt this option did not apply to their data (Figure 4).

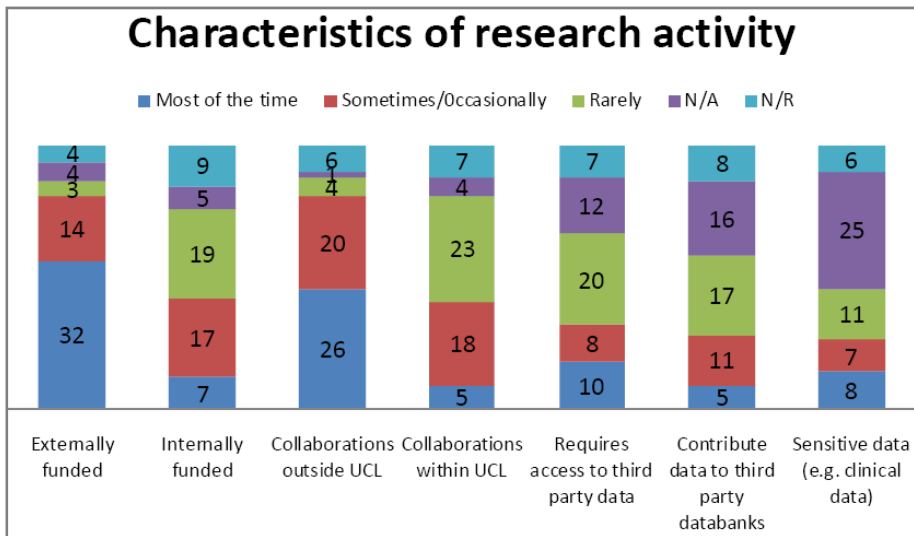


Figure 4: Characteristics of research activity.

### 7.4. Characteristics of primary research data

Following the question about the characteristics of their research activity, the participants in the survey were invited to select from a list of options those that best applied to their primary research data. In particular, they were asked whether access to their data is limited by restrictions, whether data should be retained for a

designated period after the end of the project/funding, whether their data could be useful and used by other researchers, as well as technology reliance and requirements.

The statements that the respondents rated most applicable to their data were: potential usefulness of data for researchers in other fields (31), potential long term value of data as they denoted they should be retained for a designated period after the completion of the project (31), and the importance of the contextual framework in which data were generated (25). Also, researchers appear to be conscious of the need to keep their data safe and adhere to some disaster recovery measures.

The statements that the respondents rated least relevant to their data were: primary research data can be disposed straight after the end of the project (42), access to data is limited by copyright/other restrictions (21), have requirements for use of GRID technology(24).

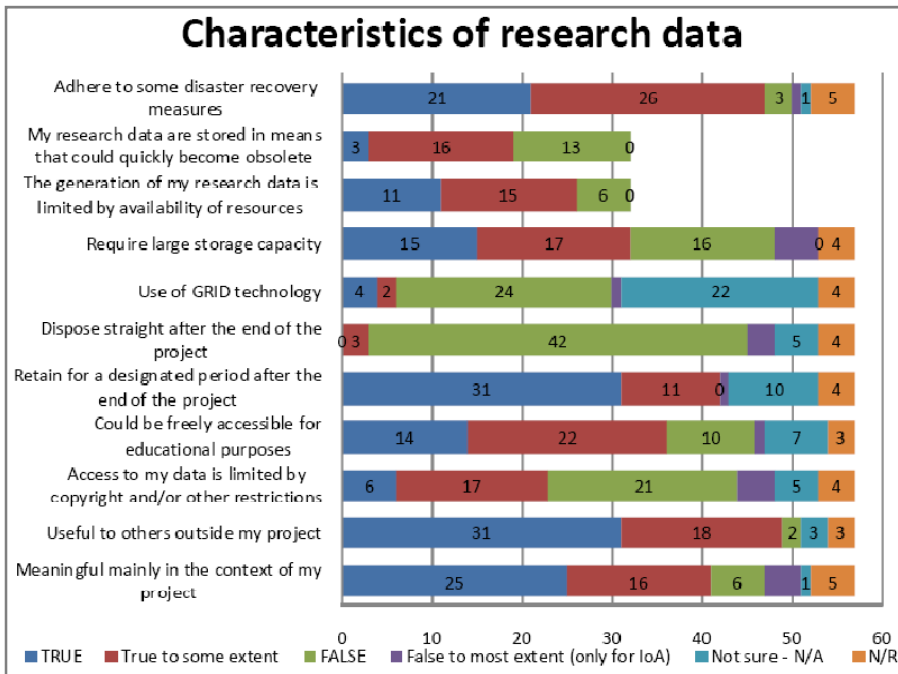


Figure 5: Characteristics of primary research data.

## 8. Discussion

Research data are complex entities. The differences in the nature of data, the level of reproducibility and the degree of processing that they have undergone are fac-

tors that can influence policy formulation (NSF, 2005). In a recent report, National Science Board members (NSF, 2005) discussed how the origins of data (computational, experimental, observation) pose different considerations for preservation and long term curation. For example, some types of observational data are unique and therefore not reproducible. Experimental conditions and costs associated with running certain experiments may act prohibitively in reproducing such data. Therefore, specific considerations may need to be applied to the preservation of such data even if those data are not unique and can be reproduced. Some computational data can be easily reproduced; however, the models and contextual framework in which data were set up, generated and processed may need to be thoroughly documented.

Given the complex nature of research data, the links and sometimes dependency on systems, services and networks, it seems that defining and documenting those links and relationships would be essential for managing and sharing such data. For example, Buchanan and Gibb (2007) write that from

an information management perspective, key data<sup>8</sup> concerns are typically associated with data protection/storage, and records management and regulatory compliance.

Previous work in this area by MacLeod and Childs (2003) report that research records comprise records of (i) the research process, (ii) the outcomes or products, (iii) the management of the process and (iv) the primary and analysed research data. The author understands this as a reference to wider information architecture comprising both records and the systems that generate, manipulate, manage and preserve them.

The preliminary results from the questionnaire confirm what one might have expected from a large research-led institution: rich and complex data; plethora of data types that originate in computational, experimental and observational research; access and use requirements that derive from the nature of interdisciplinary and collaborative research.

The examples of research data that the respondents provided demonstrate the complexity of dealing with issues around the access, use and management of such data. Perhaps the role of funders in setting requirements for preservation of research data should be further explored as staff would comply with funding regulations since the majority of the research appears to be externally funded. Furthermore, a requirement for preserving research data and/or making them ac-

---

<sup>8</sup> [As opposed to information.]

cessible to others prescribes a need for preservation planning and support. Such actions necessitate institutional acknowledgement and support.

This is further emphasised by collaborative and interdisciplinary research which is strongly encouraged at UCL. This type of research can thrive under models of data sharing (e.g. Physics and ATLAS experiment, collaborations of researchers with cultural heritage institutions, etc.). Although the participants in the survey did not specify dealing with sensitive data in detail or data which require specific access and use arrangements, it was indicated that most of the data they produce could be accessible for educational purposes. Although copyright and/or other restrictions on access to data are complex matters; it appears that in principle, it can be addressed providing that the use is for educational or non commercial research purposes.

Most of the respondents recognise the long term value of primary research data and the usefulness that they can have for other researchers. Retaining research data for a designated period after the end of the project/funding period was indicated as a requirement and common practice amongst researchers. Revisiting data produced in earlier years for inspiration and re-use to initiate further research is not uncommon.

Primary research data come in various formats and can be quantitative or qualitative. However, they do not have intrinsic value until they can be exploited. The range of values that primary research data can have was difficult to classify using the values 'minor', 'important' and 'vital' attributes suggested in the methodology. That was for various reasons: it was not always clear who owned the data (funder, researcher, university, etc.), the data may assume different values during the life-cycle of the project and after the project, the data may assume different or, in some cases, no value when they are stripped from the context in which they were generated, etc.

Therefore, the contextual framework is considered important for the meaning and value that the data assume. Perhaps the use of rich metadata can be further explored in this area, particularly if it can be conducted automatically and embedded into the research process.

## 9. Conclusions and future research

This paper presented some preliminary results from a questionnaire survey that was undertaken as part of the UCL DAF implementation project. The survey aimed at gathering information about primary research data, explore some of the issues around access to such data and factors for long term preservation.

During the planning of the audits it was evident that members of staff at all audited departments were genuinely concerned about the management of their research data and felt they had a moral responsibility for the storage, preservation and access (current and future) to the data. Scientists in some fields actively share their data. Re-use of their own data is something the academic/research staff already practise and demonstrated interest in the potential uses of their data by scholars in other fields (e.g., use of archaeological images by scholars in Humanities, use of audio visual material in Life Sciences, use of numeric data for text mining, etc.). Ownership, copyright and restrictions on access to research data can be hazy areas for academic and research staff. The brief examples that the respondents provided indicate that ownership may lie with the creator (researcher(s), etc.) the funders, the institution, and/or, in cases, be shared between them.

A full examination of these results, together with the other findings from the interviews and examples of primary research data from the UCL DAF project, is expected to enrich our understanding of the many issues around access, use, management and preservation of primary research data.

### **Acknowledgements**

The UCL DAF pilot was funded by the JISC (Joint Information Systems Committee, UK). We would like to extend our thanks to all academic and research staff who participated in our pilot and contributed with useful information and suggestions.

### **Notes and References**

- [1] JOINT INFORMATION SYSTEMS COMMITTEE (JISC). Digital Repositories programme 2007-8. Information available at <http://www.jisc.ac.uk/whatwedo/programmes/digitalrepositories2007.aspx> (March 2009).
- [2] DATA AUDIT FRAMEWORK DEVELOPMENT (DAFD). Information available at: <http://www.jisc.ac.uk/whatwedo/programmes/digitalrepositories2007/dataauditframework.aspx> (March 2009).
- [3] RLUK-RUGIT UK Research Data Service Feasibility Study. Information available at: <http://www.ukrds.ac.uk/> (March 2009).
- [4] FRY, J. et al. Identifying the benefits of curating and sharing research data, 2008. Available at [http://ie-repository.jisc.ac.uk/279/2/JISC\\_data\\_sharing\\_finalreport.pdf](http://ie-repository.jisc.ac.uk/279/2/JISC_data_sharing_finalreport.pdf) (March 2009).
- [5] SWAN, A. and SHERIDAN, B. The skills, role and career structure of data scientists and curators: an assessment of current practice and future needs, Key Perspectives Ltd, 2008. Available at <http://ie-repository.jisc.ac.uk/245/>



- (March 2009).
- [6] LYON, L. Dealing with data: roles, rights, responsibilities and relationships, 2007. UKOLN. Available at [http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing\\_with\\_data\\_report-final.pdf](http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing_with_data_report-final.pdf) (March 2009).
  - [7] NATIONAL SCIENCE BOARD. Long-lived digital data collections enabling research and education in the 21st century. Technical Report NSB-05-40, National Science Foundation, September 2005. Available from: <http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf> (March 2009).
  - [8] MACLEOD, J. and CHILDS, S. Managing primary research data and records in HE institutes, Northumbria University; Information Management Research Institute, 2003. Available at <http://www.northumbria.ac.uk/static/5007/ceispdf/resrecrep.pdf> (March 2009).
  - [9] VAN RAAN, A.F.J. Bibliometric statistical properties of the 100 largest European research universities: Prevalent scaling rules in the science system. *Journal of the American Society for Information Science and Technology*, 59(3), 2008 pp. 461-475. DOI: 10.1002/asi.20761. Pre-print available at <http://arxiv.org/abs/0704.0889> (March 2009).
  - [10] DATA AUDIT FRAMEWORK METHODOLOGY. Information available at <http://www.data-audit.eu/methodology.html> (March 2009).
  - [11] HUMANITIES ADVANCED TECHNOLOGY AND INFORMATION INSTITUTE (HATII). Information available at <http://www.hatii.arts.gla.ac.uk/> (March 2009).
  - [12] DIGITAL CURATION CENTRE (DCC). Information available at <http://www.dcc.ac.uk/> (March 2009).
  - [13] BUCHANAN, S. and GIBB, F. The information audit: role and scope. *International Journal of Information Management*, 27, 2007, pp. 159-172. doi:10.1016/j.ijinfomgt.2007.01.002.

June 2009  
Printed on demand  
by "*Nuova Cultura*"  
[www.nuovacultura.it](http://www.nuovacultura.it)

Book orders: [ordini@nuovacultura.it](mailto:ordini@nuovacultura.it)