Research Data Management in a Cultural Heritage Organisation

Tom Drysdale
Historic Royal Palaces (UK)

Abstract

Research is a core function of cultural heritage organisations. Inevitably, the undertaking of research by galleries, libraries, archives and museums (the GLAM sector) leads to the creation of vast quantities of research data. Yet despite growing recognition that research data must be managed if it is to be exploited effectively, and in spite of increasing understanding of research data management practices and needs, particularly in the higher education sector, knowledge of research data management in cultural heritage organisations remains extremely limited. This paper represents an attempt to address the limited awareness of research data management in the cultural heritage sector. It presents the results of a data management audit conducted at Historic Royal Palaces (HRP) in 2018. The study reveals that research data management at HRP is underdeveloped, while highlighting some causes for optimism. The results of the study are compared to the results of similar studies conducted in UK higher education institutions (HEIs), highlighting the many discrepancies in the ways that research data is managed at HRP and in the HE sector. Recognition of these differences and similarities, it is argued, is necessary for the development of better research data management practices and tools for the heritage sector.

Introduction

Research is a core function of cultural heritage organisations. Research into places, collections, cultures, people and events supports the delivery of exhibitions, displays, collections care and description, acquisitions, public programmes, conferences, publications and other outputs that are too many to list. But behind such outputs, and for the most part hidden from public view, is a wealth of research data – the somewhat vaguely-defined group of materials that are created, gathered or observed by researchers for the purposes of analysis and interpretation and the production of original research results.[1] In the context of the Galleries, Libraries, Archives and Museum (GLAM) sector, research data may take the form of transcripts of archival documents, oral history recordings, object records, research notes, photographs, archaeological surveys or an almost endless array of other forms. While the media, format, sources and intended uses of heritage research data may vary, however, all contribute to the interpretation, communication and preservation of heritage assets.

The management of research data is increasingly seen as an essential attribute of successful research (Pryor, 2012). Research data management involves planning, collecting, appraising, storing, preserving, governing access and re-use and transforming data. Collectively, these actions ensure that research data is valid, transparent, shareable and accessible so that it can be referenced, reused and reinterpreted by researchers in the present and in the future. The importance of good data management has been recognised by both public and professional bodies. Following suit, many funders have now made it a requirement of applicants for research grants to include data management plans in their submissions. Furthermore, the fact that most new research data is born digital makes it even more essential to manage it actively at the earliest possible opportunity.

Despite the obvious importance of research for cultural heritage organisations, little work has been done to understand the unique situation of research data management in the GLAM sector. The literature on research data management reflects the increasing level of attention paid to data curation in the academic and research sectors. Published reports and data sets from surveys conducted in higher education institutions (HEIs) in the UK and the United States reveal attitudes and approaches to research data management among academics working across different disciplines in a variety of settings (Akers and Doty, 2013; Alexogiannopoulos, Mckenney and Pickton, 2010; Cox and Williamson, 2015; Fellous-Sigrist, 2016; Kometa, 2012; Knight, 2013; Open Exeter Project Team, 2012; Parsons, Grimshaw and Williamson, 2013; Royal Veterinary College, 2013; University of Hertfordshire, 2012; University of Kent, 2015; University of Lincoln, 2012; University of Southampton, 2011; Van den Eynden, Ensom & Corti, 2013a; 2013b; Wilson, Jeffreys, Patrick, Rumsey and Jefferies, 2012). Other articles present case studies of the implementation of research data management procedures or particular facets of data management in different institutions (Guy, Donnelly and Molloy, 2013; Pryor and Donnelly, 2009; Ward, Freiman, Jones, Molloy and Snow, 2011). To date, however, little has been written about the specific characteristics of research data management in the GLAM sector.

In one of the few examples to the contrary, Edward M. Corrado and Heather Moulaison Sandy focus on the preservation of research data but fail to fully examine the

---

1   What is research data: https://www2.le.ac.uk/services/research-data/rdm/what-is-rdm/research-data

uniqueness of data management issues in the context of heritage organisations (2014). A more recent publication presents the findings of a workshop attended by practitioners of heritage research, including representatives from the British Museum, Historic England and The National Archives (Harrison, Morel, Maricevic and Penrose, 2017). The report outlines the benefits of heritage data to heritage organisations and society, including its potentially transformative value for the care, collection and curation of heritage and the ability of data to offer new ways of engaging with heritage questions. At the same time, the report also addresses some of the challenges that data poses to the heritage sector – notably the need for specialist expertise to interpret heritage data sets correctly, social uncertainty around the uses of data, and the lack of standards and infrastructure for heritage data management – and, importantly, highlights the differing needs of organisations across the sector. However, the report applies a broad definition to 'heritage data', which includes 'data related to historic buildings and environments', 'data related to… collections, geographic and provenance data, archaeological data, ecological and biodiversity data', and 'data related to specific communities, or audience related data'. While much of this 'heritage data' may also fulfil the role of 'research data', such a broad definition fails to help conceptualise the specific challenges that research data presents to heritage organisations.

It remains the case, therefore, that research data management in the heritage sector is little understood. Considering the importance of research to cultural heritage organisations – not to mention to society as a whole – uncovering the challenges and opportunities that research data management poses to the GLAM sector is essential if research data is to be made accessible and useable and safeguarded for the future. To this end, this article presents the findings of a study into research data management practices in a heritage organisation. It serves two purposes. Firstly, it provides a snapshot of research data management practices that can be used as a point of reference for other, similar organisations in the GLAM sector who wish to calibrate their own research data management frameworks. Secondly, it presents a comparison between research data management practices in a cultural heritage organisation and an aggregate of UK HEIs in order to tease out some of the similarities and differences between the GLAM and HEI sectors. Its intention is to promote discussion of the requirements of GLAMs in the field of research data management, as they differ from those of HEIs. Following a description of the methodology employed in the study, the article presents the findings of the study and discusses these with reference to data management in HEIs and their implications for research data management in the cultural heritage sector.

Methodology

A survey and a series of interviews were conducted by the author at Historic Royal Palaces in the summer of 2018. Historic Royal Palaces (HRP) is the charitable body that looks after the Tower of London, Hampton Court Palace, Kensington Palace, Kew Palace, Banqueting House and Hillsborough Castle. Its mission is to help everyone explore the story of how monarchs and people have shaped society in some of the greatest palaces ever built.[2] As well as caring for the six palaces and their estates, HRP

---

2   Historic Royal Palace – About us: https://www.hrp.org.uk/about-us. With an annual turnover of around £98m and a workforce of approximately 1,000 members of staff, HRP is broadly comparable in the size of its operation to the larger UK national museums (i.e. British Museum; Victoria and Albert Museum; Natural History Museum; Tate) and English Heritage.

also has its own collections of works of art, furniture, decorative objects, national plant collections, archaeological finds and archival material, as well as the Royal Ceremonial Dress Collection, a Designated Collection of outstanding significance. HRP's Royal Charter includes research as one of the organisation's core objectives. In 2014 HRP became an Independent Research Organisation (IRO), making it eligible to receive funding from UK Research and Innovation (UKRI) councils on the same basis as HEIs, and in 2015 received its first grant as a Lead Research Organisation under the Arts and Humanities Research Council's (AHRC) Leadership Fellow scheme.

Five active researchers were invited to participate in interviews based on their experience of research at HRP. The interviewees were chosen to represent a range of different research methods and interests. They included from HRP's Palaces and Collections Directorate a Historic Buildings Curator, a Research Curator and a Conservation Scientist, and from the organisation's Public Engagement Directorate an Interpretation Manager and a Public Engagement Producer. The interviews followed a semi-structured format, with participants responding to a series of prepared and unprepared questions on terminology, research contexts and methods, data formats, sharing, preservation and support for data management. Each interview was recorded and subsequently transcribed for further analysis. Responses were anonymised during this process.

Concurrently with the interviews, a survey was devised consisting of 25 questions (Appendix 1). Inspiration was drawn from various published surveys of data management practices in UK HEIs, with questions being tailored to the context of research data management at HRP. A link to the survey was circulated by email to all HRP employees with an invitation for those who undertake research for HRP to complete it. The survey was prefaced with a definition of 'research data', including examples, which served to assist individuals in determining whether their work constituted research or involved the processing of research data. While it was possible using this method for non-researchers to respond to the survey, since the full extent of research at HRP was not known it was decided that this was the best method for gathering responses. However, when the opportunity did present itself, known researchers – including all members of the Curatorial department – were encouraged to complete the survey.

The link to the survey was subsequently shared again via a blog posted to HRP's intranet homepage, which gave further explanation as to the purpose and intended audience of the survey, and a second email to all staff. The survey closed to new responses after three weeks on 27 July 2018. A total of 49 responses were received, of which 36 were complete and 13 were partial. None of the questions in the survey were mandatory, meaning that response rates varied for each question. In hindsight, the inclusion of compulsory questions may have garnered more representative responses, especially in the second half of the survey. While the size of the survey sample is fairly small, it does compare favourably to surveys done in several HEIs (see Table 1) and represents around 71% of active researchers at HRP according to the organisation's Annual Research Report.[3]

---

3 HRP's Annual Research Report names 69 individuals who completed research projects in the period October 2017 to September 2018. However, this number does not include those researchers who produced no research outputs within the given time period, nor does it name those who undertake (for example) market and visitor research, and is therefore used here as a guide only.

Table 1. Number of responses to RDM surveys.

	Number of responses	Response rate (where known)
Historic Royal Palaces	49	71%
University of Essex	55	13%
University of Exeter	284	–
University of Hertfordshire	67	12%
University of Lincoln	44	8%
London School of Health and Tropical Medicine	117	16%
University of Northampton	80	–
University of Nottingham	366	–
University of Oxford	314	–
Royal Veterinary College	80	23%
University of Sheffield	433	8%
University of Southampton	239	–
University College London	306	5%

There was some confusion about who the survey was aimed at, and a small number of enquiries were received from colleagues who were unsure whether their work constituted 'research'. In these cases, enquirers were provided with the definition of 'research data' used in the survey and were asked to make their own judgment about whether their roles included the creation of research data. Interestingly, it also became clear from the interviews that HRP's researchers – with the exception of those conducting science research – are not familiar with the term 'research data' and in some cases do not understand how it applies to their work. The low level of familiarity with the term 'research data' among HRP's researchers correlates with the limited understanding of the term among creative arts researchers (Guy, Donnelly and Molloy, 2013). However, as Ward et al. conclude, clear definitions are needed in order to help researchers move towards better research data management practices (Ward, Freiman, Jones, Molloy and Snow, 2011).

Results of Survey

Context and Methods of Research

Nearly three quarters of respondents (73.5%) declared Arts and Humanities as their main area of research (Figure 1). 30.6% stated that Audience or Market Research was their main area of interest, with 83.3% of the individuals conducting both kinds of research. 10.2% of respondents at HRP conduct scientific research. Small numbers of individuals also engage in social science and economic research. Respondents were able to select more than one area of research interest, and the results show that 30.6% have research interests in more than one area.
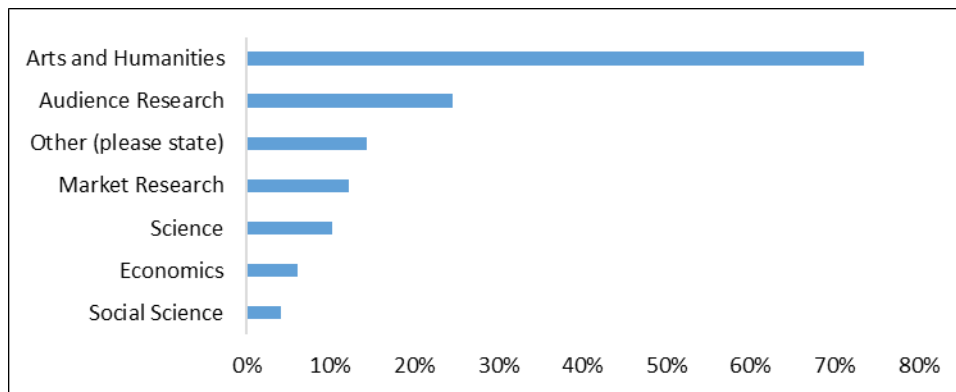
Figure 1. 'Which of the following best describes your main area of research?

Almost all HRP researchers (93.9%) create textual data. The other most common data types were identified as images (73.5%), numerical (49%) and statistical data (46.9%). HRP researchers also create bibliographic, geospatial, audio and multimedia data. 'Other' data types identified by researchers included architectural drawings, scientific data for machine-based analysis and physical data such as archaeological finds (Figure 2).
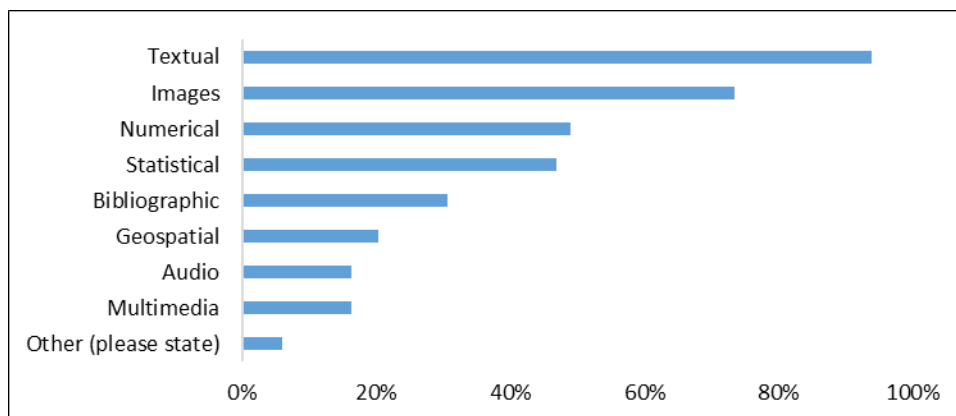

Figure 2. What kinds of data do you typically generate in the course of your research?

Correspondingly, the most common formats for research data are .doc (89.8%) and .pdf (77.6%), followed by .jpg (75.5%) and .xls (57.1%). Researchers at HRP also keep their data in .txt, .csv, .tif, .wav, .mp3, .mp4, .ppt and .html formats. 36.7% of researchers use non-digital formats such as paper. Other, more specialised formats were also identified by some researchers, including .dwg, .sav, unspecified coding and 3d file formats (Figure 3).
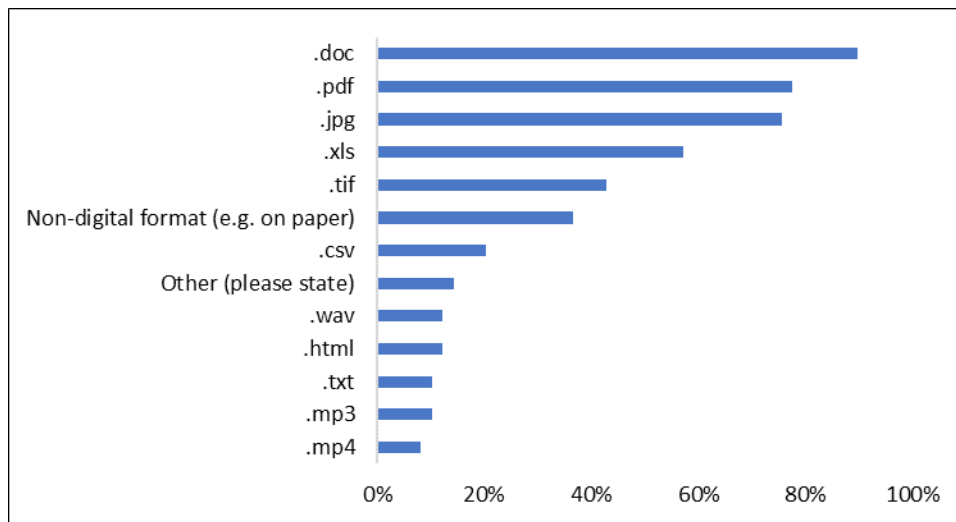
Figure 3. What formats are your research data stored in?

When asked how much data they typically create in the course of a research project, 24.5% responded that they create less than 1GB and 28.6% between 1 and 100GB. Only 12.2% of researchers at HRP create more than 100GB of data per project. Meanwhile 34.7% are not aware how much data they typically create during a project (Figure 4).
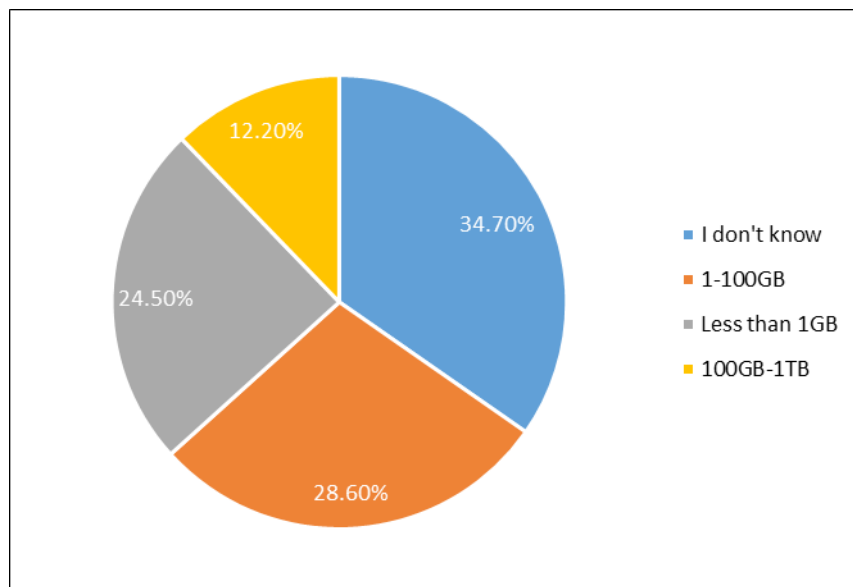
Figure 4. How much data do you typically create in the course of a research project?

Data Management Planning and Storage

The survey sought to find out about HRP researchers' experiences of working on funded projects and creating data management plans. 69.4% stated that they have never worked on a funded research project. Of the 24.5% who have done so, the most common sources of funding were the Arts and Humanities Research Council (six

respondents), Heritage Lottery Fund (three respondents) and the British Academy (two respondents). Other funders included the Economic and Social Research Council, the Modern Languages Research Association, the Leverhulme Trust, John Lyon's Charity and the European Union.

Only 10.4% of researchers at HRP have ever completed a data management plan for a research project, of which only a single respondent did so as a requirement of a funding body – namely the AHRC.

The most common location for the storage of research data by researchers at HRP is on the organisation's networked drives (91.7%). 43.8% also use portable devices, with 6.1% using portable devices but not HRP's networked drives. A quarter of respondents store research data on their personal computer or laptop, and one fifth use external hard drives. Other places where research data are stored are office computer hard drives, cloud storage and CDs/DVDs. 2% of researchers have data stored in paper files (Figure 5).
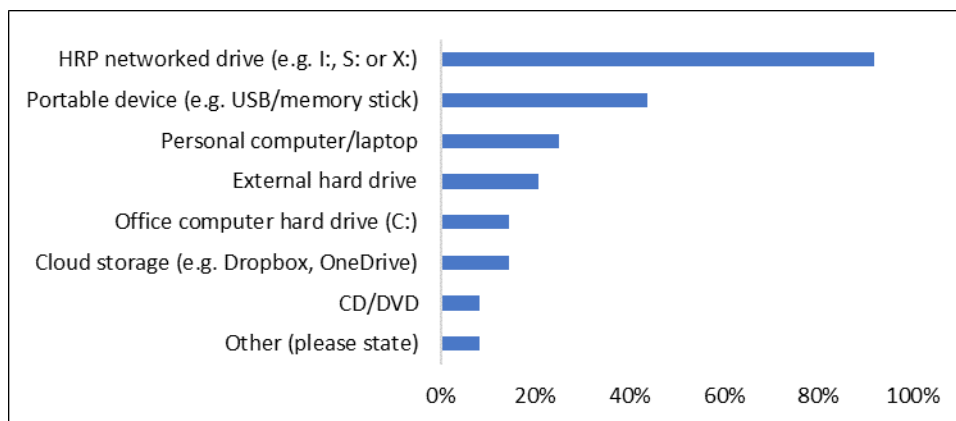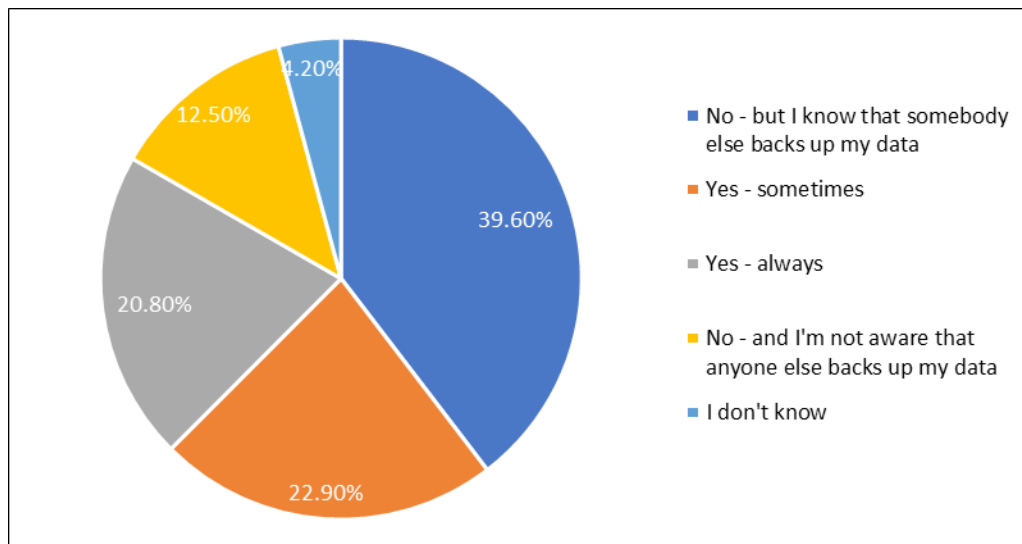

Figure 5. Where do you store your research data?

Preservation and Archiving

43.7% of researchers back up their research data 'always' or 'sometimes', while 39.6% do not back up their data but know that somebody else does. Only 12.5% do not back up their data and are not aware that anyone else does so on their behalf (Figure 6).

Figure 6. Do you actively back up your data.

51.1% of research-active individuals are 'very' or 'somewhat' confident in HRP's infrastructure for research data management. 25.5% are not very or not at all confident, and the remaining 23.4% declare themselves to be neutral.

Researchers were asked what methods they employ to ensure the integrity and validity of their research data. 59.2% responded to the question. Common answers included the use of file structures and file naming standards. None of the respondents explicitly mentioned any data validation techniques (such as data transformation or the use of checksums) or tools (such as DROID and JHOVE).

Only 32.5% of researchers said that they document or create metadata about their research data, but when asked to specify 49% acknowledged that they follow a particular standard. 87.5% of these (42.9% of all researchers) identified HRP's file naming standard.

When asked how they decide which of their research data to keep after the end of a project, 73.5% of researchers offered a response. Responses were varied, but 36.1% explicitly stated that they keep everything. 5.6% mentioned data protection as an influencing factor, and 2.8% referred to retention rules. Selection criteria include relevance, completeness, originality or uniqueness, predicted future use and accuracy of the data, but no researchers referred explicitly to defined selection policies or schedules.

Half of the respondents said that they usually need to keep their data for more than ten years and 20% said that they did not know how long they need it for. Only 5% need their data for less than one year (Figure 7).
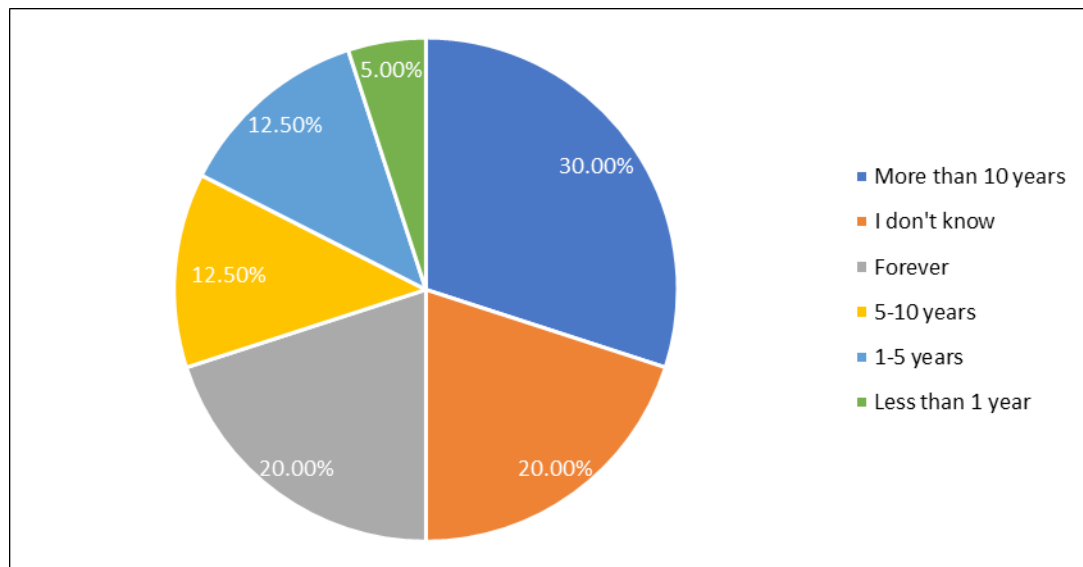
Figure 7. How long do you usually need to use, refer to or retain your research data for after the end of a project?

Access, Use and Re-use

Researchers were asked whether they ever reuse their own or others' research data for projects or purposes other than those for which they were originally created. Four fifths stated that they do reuse their own or others' research data, while 5.1% did not know.

When asked who legally owns their research data, three quarters of researchers replied 'HRP'. 10.3% believe that they alone are the owners of their research data, and another 10.3% 'don't know'. One researcher responded that the original authors of the data are the legal owners.

61.5% of researchers, 79.2% of whom are Arts and Humanities researchers, publish their research findings outside of HRP in, for example, academic journals. 33.3% of HRP researchers do not publish their findings externally, while 5.1% did not know.

Almost all HRP researchers (97.3%) are willing to share their research data with others (2.7% 'don't know'), all of whom would share their data with others within their team or others within HRP. 63.2% are happy to share their data with others in their discipline, and 47.4% with the general public. 39.5% would willingly share their research data with funders and the same percentage would share their data with publishers and with the creators of the data (Figure 8).
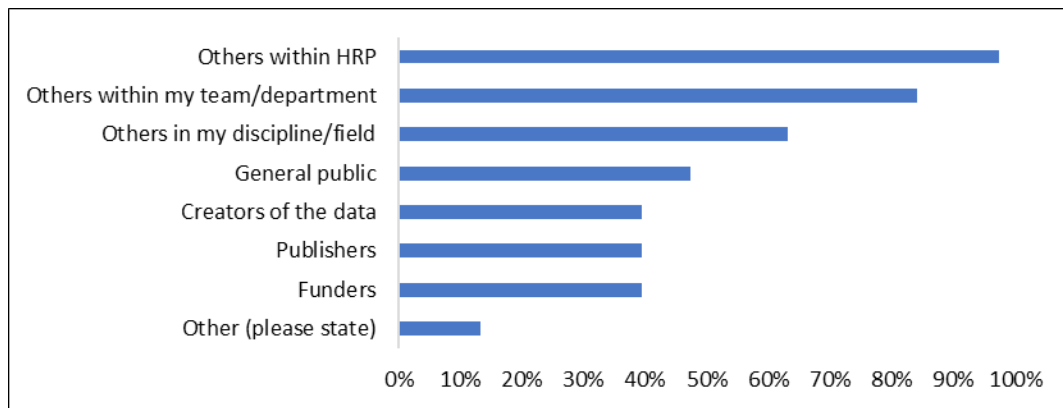
Figure 8. Who are you willing to share your research data with?

The most common obstacles to sharing data were identified as ethical issues (56.8%), lack of time or resources (54.1%) and legal issues (48.6%). 35.1% of respondents stated that commercial issues would prevent them from sharing their research data. Other barriers to sharing research data included technological restrictions, institutional policy and lack of incentives (Figure 9).
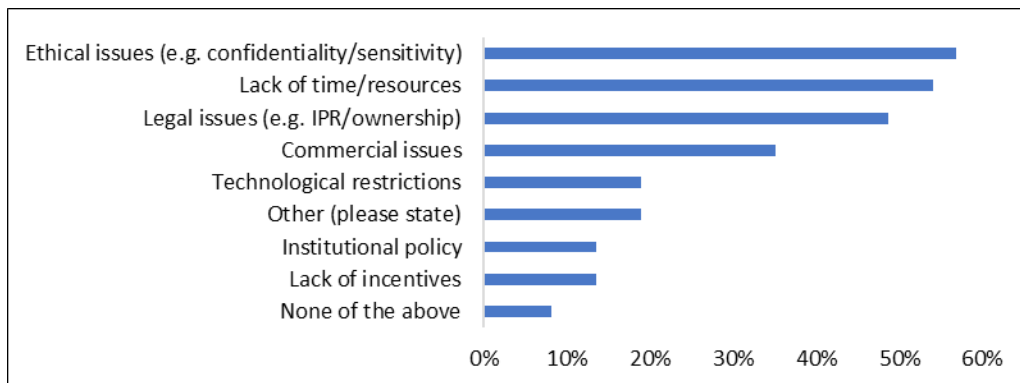
Figure 9. What prevents you from sharing your research data?

Only 5.1% of researchers at HRP have ever deposited their research data in a dedicated repository. When asked to specify which repositories they had used, researchers named the Archaeological Data Service, Oasis and the Greater London Historic Environment Record – all of which are repositories for archaeological data.

Support and Training

When asked who is responsible for the day-to-day management of their research data, 85.4% of HRP researchers responded that they themselves are responsible while 31.3% stated that responsibility lies within their team. 10.4% answered that a departmental administrator has responsibility for their research data, and 2.1% of researchers believe that IT staff are responsible, although not solely (Figure 10).
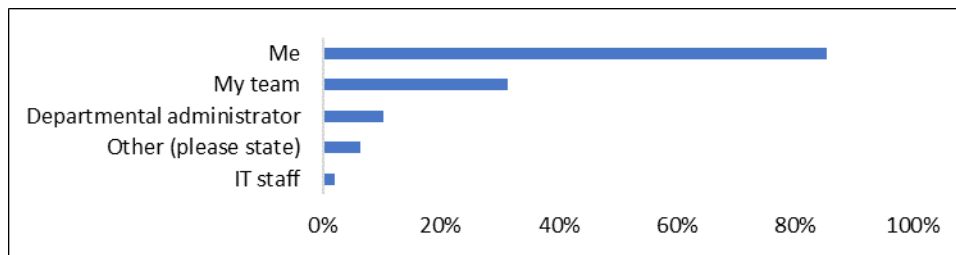
Figure 10. Who is responsible for the day-to-day management of your research data?

Around one third of respondents stated that they have received training to assist them with managing their research data against 65.8% who have not. For 61.5% of those who have had training, the training was delivered in-house. Others had received training in previous roles with different organisations, or in the course of higher education.

Only 11.8% of researchers indicated that they had no need for further training. 67.6%, however, would like training in developing a data management plan, 50% in data storage, and 44.1% in data and intellectual property rights. In addition, there is also demand for training in metadata creation, data sharing, data repositories, data formatting, data ethics and consent, Open Access and funding requirements (Figure 11).
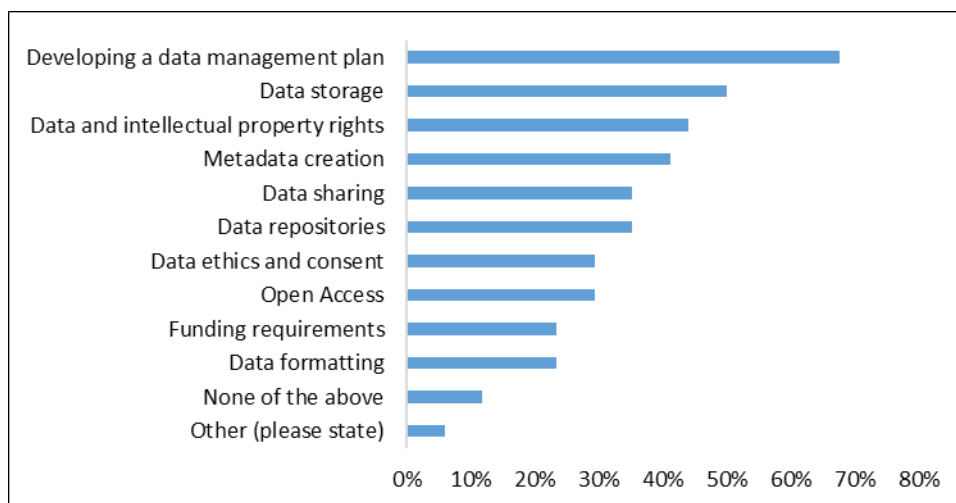
Figure 11. Which of the following areas would it be useful for you to receive training in?

Discussion

Research Data Management at HRP

Overall the results of the survey suggest that research data management at HRP is underdeveloped, while offering cause for optimism in several areas.

Predictably, most research undertaken at HRP falls under the category of 'arts and humanities', much of which is historical research. Scientists constitute a fraction of

researchers at HRP and in this context focus on heritage and conservation science. More common than science research, though, are audience and market research, although the number of arts and humanities researchers adds up to more than the total number of other researchers combined. Recognising the diversity of research undertaken at HRP is important since, as Akers and Doty have shown, research data management practices vary between disciplines (2013). Even within a relatively small-scale research institution, such as HRP, there may be differences between how research data pertaining to different disciplines is managed.

The types and formats of data used and produced by HRP researchers are all fairly common. According to Interviewee 2, a conscious effort was made to use open source and accessible software:

> 'We try to save data in an open access format, in csv format, so they are accessible for the future… For processing of the data we try to use Microsoft packages – Excel and Word – because these are the programs that our organisation is using and it's more likely we're going to use in the future.'

Another reason given for this approach was that 'we live within a heritage organisation and we can't afford' more specialist software applications. Interviewee 4 confirmed that it was not always seen as cost-efficient to purchase and install specialist software that only a small number of researchers would ever use.

Most research at HRP is funded by the organisation itself. Where research projects have received external funding, most frequently this funding has come from the AHRC. This is unsurprising, given that arts and humanities research is the main area of interest for most of HRP's researchers. While the predominance of internal funding means that researchers at HRP are less beholden to the strictures and requirements of external funding bodies, a downside is that few researchers have experience of creating data management plans for research projects. This overall lack of knowledge and awareness of the need for data management planning for research projects could be problematic for HRP, especially since many funders – including the AHRC – now require all applicants to submit data management plans to be eligible for funding.[4]

One of the positive findings of the survey was that most research data created by HRP is stored on the organisation's networked drives, which are secure, accessible and routinely backed-up. Nevertheless, a worrying number of researchers store their data on portable devices, such as USB drives, which are highly susceptible to loss, misplacement or damage. In mitigation, most researchers at HRP back-up their data or are aware that their data is backed-up somehow. However, it is important to note that backing-up data on networked drives does not constitute preservation by itself (Sabharwal, 2015). Unfortunately there is little evidence of active preservation of data at HRP, and researchers seem to have limited knowledge of data validation strategies and little awareness of data validation tools, meaning that much of HRP's research data is vulnerable and is not safeguarded in the long term.

Exacerbating this problem is the fact that there is a tendency among HRP researchers to keep all of the research data that they create during a project. In part the practice of retaining all research data stems from the lack of a formal selection policy, while HRP's institutional records retention schedule mandates the permanent retention of all records relating to research into the palaces and collections. The absence of informed selection and retention policies for research data at HRP is problematic since it

---

4 AHRC: About the new Data Management Plan: https://ahrc.ukri.org/documents/data/about-the-new-data-management-plan/

leads to hoarding or, on the other hand, the kind of ad hoc disposal that can result in the inadvertent destruction of potentially valuable data. Interviewee 4 confessed to being 'a bit of a digital hoarder' who keeps everything 'just in case', while for Interviewee 1 this was due to a 'paranoia of deleting things'. When research data are destroyed – as in the case of paper survey forms used for audience research – this is generally ad hoc (Interviewee 5). Keeping everything, meanwhile, increases the demand for storage space and consequently the cost of storage while simultaneously making the retrieval and preservation of data more difficult (Whyte and Wilson, 2010). The application of selection criteria through a retention schedule and appraisal policy for research data would help researchers at HRP to manage their data more rationally, thereby reducing the burden on HRP's resources.

Issues of preservation and storage are especially acute at HRP since only a very small number of researchers require their data for less than a year and, significantly, almost two-thirds of researchers need their data to be accessible for at least five years after the end of a project. The potential re-use of data in future projects is a major reason for keeping data indefinitely, and one that was indicated by four-fifths of HRP researchers, including Interviewee 1:

> 'I still find myself going back to things I was working on 11 years ago... I sometimes think there might be another article to be written about [an HRP building] at some point in the future so I wouldn't want to necessarily get rid of that research data and that resource I've got there.'

This long-term need for data undoubtedly has implications for the management of research data at HRP, including the need to lift preservation to a high level of priority, especially given the digital – and therefore unstable – nature of most of HRP's research data.

Re-use of research data for purposes other than those for which they were originally created is common among HRP researchers. The interviews also revealed that research data may be re-used by their original creators, by members of other teams within the organisation or by external collaborators. For example, Interviewee 5 described how 'research is our lifeblood in terms of how it is supplied by the curatorial team'. Fortunately, at HRP there is an encouraging level of clarity around intellectual property rights with regards to research data, meaning that issues around the use, sharing, publication and commercial exploitation of data should be minimal. Three quarters of researchers acknowledge that HRP is the legal owner of their research data, although anecdotally incidences of researchers retaining copies of their research files after leaving HRP employment are common. Nevertheless, this clarity around the ownership of research data is especially valuable since almost all of HRP's researchers are willing to share their research data with others. However, lack of time and resources, together with ethical issues, are major obstacles to the active sharing of data by HRP's researchers.

Comparison of HRP and HEIs

Comparing the data management practices of researchers at HRP to the practices of data management in HEIs is challenging, despite the publication of more than a dozen comparable data sets and reports from data audit surveys conducted in UK institutions. As Cox and Williamson found, there are fundamental differences between institutions that affect the ways in which data is managed (2015). It is also difficult to judge how representative of their respective institutions responses to each of the HEI surveys are.

Furthermore, the raw data is often not available for comparison. Lastly, with no standardised wording it is often difficult to make direct comparisons between responses to questions asked in different surveys. Nevertheless, the wealth of data available does mean that it is possible to make broad, qualified comparisons in certain areas, and the following discussion represents just such an attempt.

An obvious but necessary point to make is that the scale of research at HRP is much smaller than the research carried out in HEIs. HRP's c.69 research-active staff (6.9% of total staff) pales in comparison to, for example, University College London's (UCL) 11,933 research staff and students (23.7% of total staff and students) (Fellous-Sigrist, 2016).[5] Whereas institutions such as UCL consequently tend to have more developed infrastructures for the management of research data, including institutional data repositories, specialist software licences and support from qualified members of staff, HRP and other GLAM organisations do not necessarily have these same resources to hand. Far from being simply a question of resources, the presence of these tools in organisations such as UCL suggests that HEIs afford research data management a higher priority than organisations such as HRP. Heritage organisations therefore face two challenges at the outset: increasing the level of priority given to research data management, and securing the resources to implement research data management tools despite the relatively small scale of research that might be undertaken within them.

While the range of data types and formats created at HRP is fairly narrow, in HEIs it is much broader and often encompasses data that requires specialist software to collect and analyse. Moreover, for the most part HRP researchers work with small data sets of less than 100GB, and none work with volumes of data greater than 1TB. In contrast to HRP's small data sets, researchers in HEIs create bigger data sets, occasionally of over 50TB (see Table 2). The use of common data types and formats and the creation of smaller data sets at HRP means that researchers face fewer technological obstacles and that research data is less susceptible to software obsolescence or the expiry of licences than the data produced by their HEI counterparts, while the level of investment required by HRP for the storage of data will also be far smaller than it is for HEIs.

Table 2. Volume of research data created by researchers (% of researchers).

	HRP	University of Hertfordshire	University of Lincoln	University of Nottingham	RVC	University of Sheffield
<1GB	24.5	14.9	11.0	6.6	8.0	10.0
1-50GB	28.6	19.4	36.0	28.4	25.0	29.0
50-100GB		9.0	5.0	10.4	10.0	8.0
100-500GB	12.2	4.5	14.0	13.7	12.0	6.0
500GB-1TB		9.0	9.0	8.7	7.0	9.0
1-50TB	0.0	14.9	11.0	12.3	18.0	13.0
50-100TB	0.0	3.0	0.0	1.9	3.0	1.0
>100TB	0.0	4.5	0.0	1.1		1.0
Don't Know	34.7	20.9	14.0	16.9	17.0	24.0

---

5 UCL had a headcount of 38,313 students (undergraduate and postgraduate) and 11,997 staff members – including 2,687 academics and 3,479 researchers – as at 1 October 2015. UCL Student and Registry Services: https://www.ucl.ac.uk/srs/student-statistics; UCL Staff Numbers and FTE by Staff Group and Occupational Type as at 01/10/2015: https://www.ucl.ac.uk/human-resources/sites/human-resources/files/staff_numbers_oct_2015.pdf.

One area where HRP compares favourably to HEIs is in the storage of research data. Whereas HRP researchers mostly store their research data on the organisation's networked drives, researchers in HEIs tend to store their data locally on either PC or laptop hard drives, portable media such as USB drives, or in Cloud storage, rather than shared servers (see Table 3). While each of the institutions listed in Table 3 provides researchers with secure, backed-up network space for the storage of research files, the preference of HEI researchers for local storage media means that their data is more susceptible to loss or damage, since local media tends not to be routinely backed-up and may otherwise be lost, damaged or corrupted. In addition, the low percentages of researchers who use their respective repositories' networked servers for storage suggests that much of the research data produced by HEI researchers remains vulnerable.[6] On the other hand, HEI researchers are more active in the preservation of their research data and on average a greater proportion create metadata about their research data than researchers at HRP (see Table 4).

Table 3. Most common storage locations for current research data.

HRP	UCL	University of Hertfordshire	University of Kent	University of Lincoln	University of Sheffield
1. HRP networked drive	1. Hard drive of personal PC/laptop	1. Hard disk drive of PC/laptop	1. PC/laptop	1. Hard disk drive of PC/laptop	1. Portable storage device
2. Portable device	2. Hard drive of UCL PC/laptop	2. External hard drive	2. Local movable/portable media	2. External hard drive	2. Local disk
3. Personal computer/laptop	3. Cloud service	3. On shared drive/local server	3. Commercial 'free' offsite/the Cloud	3. Memory stick/USB/flash drive	3. Web-based service

Table 4. Percentage of researchers who create metadata about their research data.

	% of researchers who create metadata
Historic Royal Palaces	32.5%
University of Hertfordshire	61.2%
University of Lincoln	64.0%
University of Nottingham	59.3%
Royal Veterinary College	29.0%

---

6   Percentage of researchers who store data on institutional servers or networked drives: UCL = 26.8%; Kent = 11.8% (school/department server)/5.5% (central university server); Lincoln = 11%; Sheffield = 30% (university drive)/21% (university Google drive).

Table 5. Most common responses to the question of 'who owns your research data?' (% of respondents).

HRP	University of Hertfordshire	University of Lincoln	University of Nottingham	Royal Veterinary College	University of Southampton
1. HRP (74.4%)	1. University of Hertfordshire (59.7%)	1. You [the researcher] (73.0%)	1. I don't know (23.0%)	1. The College (64.0%)	1. Me (33.0%)
=2. Me alone (10.3%)	2. You [the researcher] (43.3%)	2. The University of Lincoln (39.0%)	2. Me (21.0%)	=2. Me (c.32.0%)	2. My school/institution (26.2%)
=2. I don't know (10.3%)	3. Your research group or cohort (37.3%)	3. Your research group or cohort (34.0%)	3. My research group (19.0%)	=2. My funder (c.32.0%)	3. I don't know (13.1%)
				=2. I don't know (c.32.0%)	

Another advantage that HRP has over HEIs is in the ownership of data. The clarity around the ownership of research data shown by researchers at HRP is not reflected in HEIs, where by contrast this is a much greyer area (see Table 5). The terms of HRP staff contracts include the assignment of all intellectual property rights to the products of their work – including research conducted during the course of employment – to the organisation. In HEIs, however, researchers commonly believe that they themselves own the research data that they create, even though this may not always be the case. For example, the University of Hertfordshire and the University of Lincoln both have clear policies stating that the institutions own the intellectual property rights to research created by staff and students.[7] Part of the complication no doubt stems from the distinction between research staff and research students, and the fact that students will usually own the rights to their own research data unless their research is funded by the university, carried out as part of a team which includes university staff, or is collected using equipment supplied by the university.[8]

Linked to the issue of ownership is the issue of data sharing. Comparison of HRP and HEIs shows that researchers in HEIs are more protective of their data than researchers who work for HRP. Whereas almost all of HRP's researchers are willing to share their research data with others, at the University of Exeter only 64.4% of researchers share their data, while 30% of researchers at the University of Lincoln say that they typically don't share their data (Open Exeter Project Team, 2012; University of Lincoln, 2012). The numbers are higher at the University of Oxford, where 69.4%

---

[7] University of Hertfordshire, Intellectual Property: https://www.herts.ac.uk/__data/assets/pdf_file/0018/232524/CA04-Intellectual-Property.pdf; University of Lincoln, Code of Practice for Research: https://cpb-eu-w2.wpmucdn.com/blogs.lincoln.ac.uk/dist/8/8024/files/2013/08/Research-Code-of-Practice-1g5523o.pdf

[8] See, for example, the University of Southampton's Intellectual Property Regulations: http://www.calendar.soton.ac.uk/sectionIV/ipr.html

would share some or all of their data without restrictions, and at the University of Southampton, where 73.8% would allow others to access their data after a project has ended (Wilson, Jeffreys, Patrick, Rumsey and Jefferies, 2012; University of Southampton, 2011). However, these figures are still lower than HRP's 97.3%. The figures appear to support Interviewee 1's claim that

> '[in HEIs] there is a real sort of territorial protection of your own subject area and the data that underpins your own subject area. If you've found a new archival source that uncovers something about your subject you're reluctant to give other people the reference to that source just in case they get there first and get to publish it first, whereas I think there's a cultural difference with museums and heritage where our sort of *raison d'être* is to share the information that we find, to share the stories of the places we work in or the collections we look after, so I think there is much more of a sharing culture in the way that [heritage organisations] do research.'

However, it is important to acknowledge that researchers in HEIs who conduct sensitive research – for example, involving human subjects – often have good reason and may be bound by ethics committees and protocols to restrict access and afford greater protection to their data. Furthermore, while it might be true that HRP's researchers are more willing to share their research data with colleagues, in practice it is HEI researchers who are more proactive in their sharing of data with the general public through data repositories. Only 5.1% of HRP researchers have ever deposited data in a data repository, whereas 45% of researchers at the University of Essex have done so, and 57.9% at Newcastle, 56% at Northampton and 30% at the RVC have or would use a public or institutional repository (Van den Eynden, Ensom and Corti, 2013b; Kometa, 2012; Alexogiannopoulos, McKenney and Pickton, 2010; Royal Veterinary College, 2013).

One area of common ground between HRP and HEI researchers is in demand for training. Both sets of researchers typically seek training in the same areas, namely data management planning, data storage and intellectual property rights (see Table 6). Training in digital curation remains a contested issue, and questions remain around who requires training and how this training should be delivered and accredited (Pryor and Donnelly, 2009). However, the strong correlation between training needs in HRP and in HEIs suggests that shared solutions may be available.

Table 6. Most common responses to which areas of training would be useful.

HRP	University of Essex	University of Nottingham	Royal Veterinary College	University of Sheffield	UCL
1. Developing a data management plan	1. Planning for data management and sharing	1. Developing a research data management plan	1. Developing a research data management plan	1. Developing a research data management plan	1. Storage and preservation of data (not personal or sensitive)
2. Data storage	2. Storage and back-up of data and files	2. Storing your data	2. The use of data repositories and Open Access	=2. Storing your research data	2. Data management plans
3. Data and intellectual property rights	3. Ethical and legal aspects of data sharing and re-use	3. Creating metadata for data	3. Copyright and Intellectual Property Rights	=2. Copyright and Intellectual Property	3. Storage and preservation of data (mostly personal or sensitive)

Conclusion

The survey and interviews conducted for this study have shown that there is a general recognition among HRP's researchers of the need to manage research data but that on a practical level research data management is underdeveloped. Research at HRP is characterised by a high proportion of arts and humanities research, resulting in the creation of relatively small quantities of data in largely standard formats. There is a lack of awareness among researchers of the technical aspects of research data management, coupled with little understanding of data management planning requirements and a lack of thought regarding the long-term preservation of research data. However, HRP researchers demonstrated a high level of personal responsibility and an encouraging use of existing infrastructure (where it does exist), as well as a willingness to share data and an appetite for further training and instruction.

In comparison with representatives of the HE sector, HRP's research data presents less variety in terms of formats and lower technical demands (Table 7). HRP researchers also seem to work in a more homogenous way, with a greater reliance on internal systems and more recognition of individual rights and responsibilities concerning research data. Researchers at HRP have less experience of funders and their requirements but are open to sharing and re-using their research data, despite a lack of tools and knowledge to do so effectively. The priority areas for training – namely data management planning, sharing and IPR – are shared by HRP and HEIs.

Table 7. Comparison of research data management practices in HRP and HEIs.

HRP	HEIs
Small data sets	Occasionally large data sets
Few different data types and formats	Wide variety of data types and formats
Centralised storage and backing-up	Distributed/local storage and ad hoc backing-up
Personal responsibility for RDM with little support	Support of Research Data Officers/ Managers/ Librarians
Clarity over ownership of data	Uncertainty over ownership of data
Little experience of funded research and data management planning	Considerable experience of funded research and data management planning
Openness to sharing data with colleagues	Openness to sharing data with the public
Very little use of data repositories	Some use of data repositories
Main obstacles to sharing are ethical and commercial issues and lack of time/resources	Main obstacles to sharing are sensitivity and data subject privacy
Training sought for data management planning, data storage and IPR	

Clearly, then, there are differences and similarities between the ways that research data is managed in HRP and in the HE sector. Recognition of these differences is necessary for the development of the bespoke solutions that are required if organisations such as HRP are to manage their research data according to best practices. On the other hand, knowledge of the similarities presents the opportunity for organisations like HRP to learn by example and benefit from the existing policies, practices and procedures already developed by and for HEIs. The potential for collaboration is especially pertinent for those heritage organisations that have existing partnerships in place with HEIs. While further studies in cultural heritage organisations must be done before any generalisations can be made about the differences in research data management between the HE and heritage sectors, this study nevertheless serves to indicate the types of conclusions that such studies might be expected to reach and draws attention to the unique requirements of research data management in a typical UK heritage organisation.

Acknowledgements

References

Akers, K. G., & Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation 8*(2). doi:10.2218/ijdc.v8i2.263

Alexogiannopoulos, E., McKenney, S., & Pickton, M. (2010). *Research data management project: A DAF investigation of research data management practices at The University of Northampton.* Retrieved from http://nectar.northampton.ac.uk/2736/1/Alexogiannopoulos20102736.pdf

Corrado, E.M., & Moulaison Sandy, H. (2014). *Digital preservation for libraries, archives and museums*. Lanham, MD: Rowman & Littlefield.

Cox, A., & Williamson, L. (2015). The 2014 DAF survey at the University of Sheffield. *International Journal of Digital Curation 10*(1). doi:10.2218/ijdc.v10i1.362

Fellous-Sigrist, M. (2016). *UCL researchers and their research data: Practices, challenges & recommendations*. Retrieved from http://discovery.ucl.ac.uk/1540140/

Guy, M., Donnelly, M., & Molloy, M. (2013). Pinning it down: Towards a practical definition of 'research data' for creative arts institutions. *International Journal of Digital Curation 8*(2). doi:10.2218/ijdc.v8i2.275

Harrison, R., Morel, H., Maricevic, M., & Penrose, S. (2017). *Heritage and data: Opportunities and challenges for the heritage sector*. London: British Library. Retrieved from https://heritage-research.org/2017/10/23/heritage-data-report

Kometa, S. (2012). *Research data management requirements online survey report*. Retrieved from https://research.ncl.ac.uk/media/sites/researchwebsites/iridium/iridium_online_survey_report_17_8_2012_v2.1_SK.pdf

Knight, G. (2013). *Research data management at London School of Hygiene and Tropical Medicine: Web survey report.* Retrieved from http://blogs.lshtm.ac.uk/rdmss/files/2013/04/LSHTM-RDM-Web-Survey-Report.pdf

Muir, A. (2006). *Preservation, access and intellectual property rights challenges for libraries in the digital environment*. London: Institute for Public Policy Research. Retrieved from https://www.ippr.org/publications/preservation-access-and-intellectual-property-rights-challenges-for-libraries-in-the-digital-environment

Nassiri, S., & Worthington, B. (2012). *Digital Asset Framework (DAF) survey analysis.* Retrieved from http://research-data-toolkit.herts.ac.uk/document/rdtk-data-asset-survey-digest-july-2012/

Open Exeter Project Team. (2012). *Summary findings of Open Exeter DAF survey.* Retrieved from http://hdl.handle.net/10036/3689

Parsons, T., Grimshaw, S., & Williamson, L. (2013). *Research data management survey, the University of Nottingham.* Retrieved from http://admire.jiscinvolve.org/wp/files/2013/02/ADMIRe-Survey-Results-and-Analysis-2013.pdf

Pryor, G. (ed.) (2012). *Managing Research Data*. London: Facet Publishing.

Pryor, G., & Donnelly, M. (2009). Skilling up to do data: Whose role, whose responsibility, whose career? *International Journal of Digital Curation 4*(2). doi:10.2218/ijdc.v4i2.105

Royal Veterinary College. (2013). Analysis of research data management survey. Retrieved from https://www.rvc.ac.uk/Media/Default/Research%20Data %20Management/RDM%20Survey%20Analysis%20v1.0-1.pdf

Sabharwal, A. (2015). *Digital curation in the digital humanities. Preserving and promoting archival and special collections*. Kidlington: Chandos Publishing.

University of Hertfordshire. (2012). *Research data assets survey*. Retrieved from http://research-data-toolkit.herts.ac.uk/document/rdtk-data-asset-survey-results-jul-2012/

University of Lincoln. (2012). *Research data management survey.* Retrieved from http://orbital.blogs.lincoln.ac.uk/files/2012/04/DAF-summary-anonymised.pdf

University of Southampton. (2011). *Data management questionnaire results: IDMB Project.* Retrieved from http://eprints.soton.ac.uk/196243/1/IDMB_Survey_Report.pdf

Van den Eynden, V., Ensom, T., & Corti, L. (2013a). *University of Essex staff research data survey*. UK Data Archive, University of Essex. Retrieved from https://dam.data-archive.ac.uk/reports/research/rdessex_staffsurveyreport.pdf

Van den Eynden, V., Ensom, T., & Corti, L. (2013b). *University of Essex staff research data survey results*. UK Data Archive, University of Essex. Retrieved from https://dam.data-archive.ac.uk/reports/research/rdessex_staffsurveyresults.pdf

Ward, C., Freiman, L., Jones, S., Molloy, L., & Snow, K. (2011). Making sense: Talking data management with researchers. *International Journal of Digital Curation 6*(2). doi:10.2218/ijdc.v6i2.202

Whyte, A., & Wilson, A., (2010). *How to appraise and select research data for curation*. Retrieved from http://www.dcc.ac.uk/sites/default/files/documents/How%20to %20Appraise%20and%20Select%20Research%20Data.pdf

Wilson, J., Jeffreys, P., Patrick, M., Rumsey, S., & Jefferies, N. (2012). *Results of the 2012 University of Oxford research data management survey*. Oxford University. Retrieved from http://damaro.oucs.ox.ac.uk/docs/OxfordRDMsurvey2012_public.xlsx

Appendix 1: HRP Research Data Management Survey

This survey is designed to find out about how you use, manage and preserve your research data.

Research data is defined as 'data that is collected, observed or created for purposes of analysis to produce original research results'. Examples might include survey responses, transcripts or photographs. Research data do not generally include published outputs, such as articles, lectures and reports.

About your research data

1. Which of the following best describes your main area(s) of research? (Select all that apply)

- Arts and Humanities

- Science

- Social Science

- Economics

- Market Research

- Audience Research

- Other (please state) _____

2. What kinds of data do you typically generate in the course of your research? (Select all that apply)

- Textual

- Numerical

- Statistical

- Geospatial

- Images

- Audio

- Multimedia

- Bibliographic

- Other (please state) _____

3. What file formats are your research data stored in? (Select all that apply)

- .doc

- .pdf
- .txt
- .xls
- .csv
- .jpg
- .tif
- .wav
- .mp3
- .mp4
- .html
- .dta
- Non-digital format (e.g. on paper)
- Other (please state) _____

4. How much data do you typically create in the course of a research project?

- Less than 1GB
- 1-100GB
- 100GB-1TB
- More than 1TB
- I don't know

5. Have you ever worked on a funded research project (including current projects)?

- Yes
- No
- I don't know

5a. If yes, who funded the project?

_____

Data management and storage

6. Who is responsible for the day-to-day management of your research data? (Select all that apply)

- Me
- My team
- Departmental administrator

- IT staff

- Other (please state) _____

- I don't know

7. Have you ever completed a data management plan for a research project?

- Yes

- No

- I don't know

7a. If yes, was this a requirement of a funding body?

- Yes (please name) _____

- No

- I don't know

- Not applicable

8. Where do you store your research data? (Select all that apply)

- Office computer hard drive (C:)

- HRP networked drive (e.g. I:, S: or X:)

- External hard drive

- Portable device (e.g. USB/memory stick)

- CD/DVD

- Cloud storage (e.g. Dropbox, OneDrive)

- Personal computer/laptop

- Other (please state) _____

9. Do you actively back up your data?

- Yes – always

- Yes – sometimes

- No – but I know that somebody else backs up my data

- No – and I'm not aware that anyone else backs up my data

10. How confident are you that HRP's infrastructure meets your needs for research data management?

- Very confident

- Quite confident

- Not very confident

- Not at all confident

Preservation and archiving

11. What methods do you use to ensure the integrity and validity of your digital research data? (e.g. file structures; encryption; documentation; migration/transformation; validation tools)

_____

12. Do you ever document or create metadata about your data?

- Yes

- No

- I don't know

12a. If yes, do you follow any particular standards? (e.g. HRP's file naming standard)

- Yes (please specify) _____

- No

- I don't know

13. How do you decide which of your research data to keep after the end of a project?

_____

14. How long do you usually need to use, refer to or retain your research data for after the end of a project?

- Less than 1 year

- 1-5 years

- 5-10 years

- More than 10 years

- Forever

- I don't know

15. Do you ever re-use your own or others' research data for projects or purposes other than those for which they were originally created?

- Yes

- No

- I don't know

- Not applicable

Depositing and sharing

16. Who legally owns the research data you create?

- Me alone

- HRP

- Third party e.g. university/collaborator

- My funder

- I don't know

- Other (please state) _____

17. Do you publish your research findings outside of HRP (for example, in academic journals, at conferences etc.)?

- Yes

- No

- I don't know

18. Are you willing to share your research data with others?

- Yes

- No

- I don't know

18a. If yes, who are you willing to share your research data with? (Select all that apply)

- Others within my team/department

- Others within HRP

- Others in my discipline/field

- Funders

- Publishers

- General public

- Creators of the data

- Other (please state) _____

19. What prevents you from sharing your research data? (Select all that apply)

- Ethical issues (e.g. confidentiality/sensitivity)

- Legal issues (e.g. IPR/ownership)
- Commercial issues
- Technological restrictions
- Institutional policy
- Lack of time/resources
- Lack of incentives
- None of the above
- Other (please state) _____

20. Have you ever deposited your research data in a dedicated repository?

- Yes (please specify) _____
- No
- I don't know

Support and training

21. Have you ever received any training to assist you with managing your research data?

- Yes (please describe) _____
- No
- I don't know

22. Which of the following areas would it be useful for you to receive training in? (Select all that apply)

- Developing a data management plan
- Data storage
- Data sharing
- Data formatting
- Metadata creation
- Data ethics and consent
- Funding requirements
- Data and intellectual property rights
- Data repositories
- Open Access
- None of the above
- Other (please state) _____

23. Please use this space for any additional comments about your research data management practices or needs.

—————————————————————

24. If you would be willing to be contacted about your responses to this survey, please provide your email address in the box below.

—————————————————————