



Project Information			
Project Identifier			
Project Title	Service Oriented Toolkit for Research Data Management		
Project Hashtag	#rdtk_herts		
Start Date	03/10/2011	End Date	31/06/2013
Lead Institution	University of Hertfordshire		
Project Director	Dr David Ford		
Project Manager	Dr William Worthington		
Contact email	w.j.worthington@herts.ac.uk ; s.hajnassiri2@herts.ac.uk		
Partner Institutions	Digital Curation Centre		
Project Webpage URL	http://www.herts.ac.uk/research-data-toolkit		
Programme Name	Managing Research Data		
Programme Manager	Dr Simon Hodson		

Document Information			
Author(s)	Sara Hajassiri, Bill Worthington		
Project Role(s)	Project analyst, Project Manager		
Date	24/06/07	Filename	RDTK project benefit analysis
URL			
Access	Version 2		

1 Introduction

Over the last 18 months and through the JISC funded RDTK project, the Information Hertfordshire has managed to develop thorough understanding of the research data scope and management requirements within the university. The project team also established the foundation for the required process and people change through involvement in developing and amending policies and procedures, developing, pilot test of a series of new infrastructure and IT solution, and developing and delivering of a series of trainings in support of good practice research data management.

This report aims to provide evidence of the benefits delivered to the UH research community and supporting units through the RDTK project. The identified benefits can be classified to support the institution in the following areas:

- Raised awareness of RDM/ Improved data management skills
- Improved institutional understanding of RDM requirements and providing support
- More efficient management of working data
- More data available for re-use
- Improved compliance With Local DM policy and funder requirements

2 Established Benefits

2.1 Raised awareness of RDM/ Improved data management skills

The project team have arranged a series of workshops and project events to spread the words about the importance of research data management, scope of the project and the services and supports that the Information Hertfordshire is offering to help in managing research data. The awareness sessions and trainings can be divided into four categories:

- Stakeholder forum meetings
- Lunch time seminars
- Presenting to the individual research groups through requesting agenda in their research group gathering
- Training programme for students and staff

Table 1 shows the number of attendees in each event, and where applicable their feedback about content of the session. Feedback was collected in all sessions but stakeholder forum.

Date	Program	Title	Aud.	No. Att.	Feedback											
					a	b	c	d	e	f	g	h	i	j	k	av.
22/03/12	Stakeholder Forum Meeting	RDM update	Staff	16												
27/03/12	Lunch time Seminar	RDM update	Staff	12	4.2	3.9	3.6	4.4	4.0	3.9	3.9	3.8	NA	3.6	4.0	3.9
30/04/13	Staff Dev.	RDM update	Staff	6	4.0	4.0	3.4	4.2	4.2	4.0	4.0	4.0	4.0	4.4	4.0	4.0
12/06/12	Centre 4 regional & local history research	RDM update	Staff	9												
13/06/12	PAM School meeting	RDM update	Staff	20												
03/07/12	H&HSRI Research group	RDM update	Staff	12												
10/12/12	Stakeholder Forum Meeting	RDM update	Staff	16												
16/10/12	Lunch time Seminar	RDM update	Staff	16	4.3	3.6	3.7	4.4	3.8	3.7	3.7	3.8	3.3	3.9	4.0	3.8
23/10/12	Graduate student training	Intro. RDM	PGRS	4	4.3	4.3	4.3	4.3	4.5	4.5	3.8	4.0	4.0	4.5	4.5	4.3
04/12/12	Staff Dev.	Encryption	Staff	22	3.9	4.0	3.7						4.0	4.1	4.1	3.9
22/01/13	Staff Dev.	Encryption	Staff	14	3.9	4.3	4.1						4.1	4.3	4.3	4.2
13/05/13	Graduate student training	Intro. RDM	PGRS		3.9	4.1	3.6	4.5	3.8	3.6	3.6	3.0	4.3	4.0	4.0	3.9
17/05/13	Staff Dev.	RDM update	Staff	14	4.7	4.7	4.7	5	4.3	4.3	4.3	3.5	4.3	5.0	5.0	4.6

Table 1: Summary of the feedback from training sessions throughout UH

The questions asked refer to a) the content, b) the length of session, c) the UH infrastructure, d) DMPonline, e) funder requirements, f) UH storage, g) sharing solutions, h) archiving, i) encryption,

j) the document management system (DMS), and k) the overall experience. Each was scored out of 5 where 1 strongly disagreed and 5 strongly agree with the statement.

Throughout the lifetime of project, the RDTK team has continuously provided UH and wider research community and support units with update about the project progress and findings through the [project blog](#):

Table 2 shows the statistics related to the page view of the blog's posts.

Page name/ creation month	Pageviews	Avg. Time on Page
/2011/10/hello-world/	18	108.27
/2011/10/uh-data-management-policy-and-guide/	12	65.63
/2011/11/mrd-programme-launch-poster/	25	12.33
/2011/11/progress-report-month-2-november-2011/	16	283.13
/2011/11/rdtk_herts-progress-report-month-1-october-2011/	6	413.50
/2011/11/reflection-on-rdmf7-and-dcc-cambridge/	5	771.00
/2011/12/jiscmr-d-2011-2013-launch-meeting---blog-analytics/	5	92.75
/2011/12/jiscmr-d-2011-2013-launch-meeting---institutional-approach-breakout-group/	3	163.00
/2011/12/jiscmr-d-2011-2013-launch-meeting---thematic-session-on-business-case/	19	54.67
/2011/12/jiscmr-d-2011-2013-launch-meeting-biomedicalhealth-breakout-group/	7	483.50
/2012/01/data-management-planning-using-dmponline/	17	260.22
/2012/01/networked-secure-storage-faq/	265	98.06
/2012/01/progress-report-month-3-december-2011/	9	45.00
/2012/01/u-drive-and-beyond/	9	15.00
/2012/01/what-happened-next/	18	113.27
/2012/02/information-hertfordshire-lunchtime-presentation-on-rdm/	41	73.90
/2012/02/progress-report-month-4-january-2012/	8	569.50
/2012/02/uh-research-data-policy-in-a-nutshell/	133	185.19
/2012/03/progress-report-month-6-march-2011/	108	103.25
/2012/03/reflections-on-jiscmr-dcc-policy-workshop/	36	363.76
/2012/04/steering-group-december-2011/	4	172.00
/2012/05/atira-pure-cris-roundup/	218	153.23
/2012/05/research-data-assets-survey/	45	129.69
/2012/05/steering-group-march-2011/	16	153.45
/2012/06/document-management-for-clinical-trials/	381	97.85
/2012/06/rdm-audit-and-project-benefit-metrics/	115	171.86
/2012/08/data-asset->survey-results/	1	7.00
/2012/08/data-asset-survey-results/	88	161.68
/2012/08/rdmtpa/	92	142.23

/2012/10/data-encryption/	119	210.21
/2012/10/does-for-datasets/	219	198.95
/2012/10/files-in-the-cloud/	61	248.21
/2012/10/files-in-the-cloud/www.hrc3.org.uk/	2	186.50
/2012/10/jiscmr-d-programme-progress-workshop2012/	134	151.64
/2012/12/progress-report-month-15-dec-2012/	47	149.35
/2012/12/progress-report-month-15-march-2011/' rel='external nofollow	1	5.00
/2013/04/comparison-of-open-licenses/	44	410.06
/2013/04/rdtk-jiscmr-d-programme-workshop-aston2013/	48	99.08
/2013/04/research-project-file-plan/	21	135.77
/2013/04/zendto-file-exchange-in-the-cloud/	22	62.08
/2013/05/test-blog/	2	0.00
/2013/05/the-cost-of-a-bit-of-a-ddud/	62	492.67
Total	2502	157.07

Table 2 the RDTK project blog page views

2.2 Improved institutional understanding of RDM requirements and provision of support

One of the main aims of the project was to develop a better understanding of research data management practices and requirements within UH. For this purpose first a series of interviews was carried and case studies produced, these are attached in annex I. Also a survey of research data scope and managing practices was carried out with a response rate of %12.5. The result of the survey in particular was helpful to give us an estimation of the volume of research data in the university, the average storage requirement and the requirements for training and staff development.

The university has also introduced expanded the existing Electronic Document and Records Management systems (EDRMS) to support managing all documents related to a research project. After trial of the system with Centre for Lifespan and Chronic Illness Research (CLiCIR) for a drug trial project, and consulting numerous researchers and professional staff across the university, a file plan structure was developed, which is now using for structuring different documents related to a research project.

The system is now being used by 9 research projects and more than 40 researchers.

	Research group	Research project	Internal users	External Collaborators
1	MIRU	University Choice	1	3
2	CRIPPAC	ASSERTID	2	0
3	CRIPACC	BPSD	2	0
4	CRIPPAC	CoDem	3	3
5	CLICR	EDS	2	0
6	CRIPPAC	HBSC	5	0
7	CRIPPAC	OPTIMAL	4	7
8	CRIPPAC	PharmchemJLS	5	0
9	CRIPPAC	RAPPORT	4	1

Table 3 Current users of EDRMS

2.3 More efficient management of working data

The Electronic Document and Records Management systems (EDRMS) and its potential to answer the needs of Research Data Management is now recognized by the UH research community. Table 3 shows that currently 9 research project are using the system and we are receiving requests from more research project to access the system.

The increasing use of the system clearly demonstrate a shift within UH research community towards efficient management of their research documents and data.

Another example of adopting more efficient data management practices among researchers in the UH is the interest towards using Truecrypt software for encryption of sensitive data. As Table 1 shows, so far 36 people have attended the Data Encryption training. We have also had 1-1 trainings for some users. Following is a quote from a user about the financial benefits of using free encryption software in comparison of PGP encryption:

“We were looking at PGP encryption but it looks like Truecrypt will save us £110 per laptop”.

2.4 More data available for re-use

So far we have accepted 6 datasets to be deposited in the UH Research Archive (UHRA). Details are shown in Table 4.

Research group	Data	Format	Volume	Metadata
Local History	71 interviews	WAV + MP3	28GB	1 Excel file metadata
Local History	40 interviews	WAV + MP3	4GB	1 Excel file metadata

Local History	59 interviews	WAV + MP3	9GB	1 Excel file metadata
Local History	139 interviews	WAV + MP3	64GB	1 Excel file metadata
Liesure Studies	220 questionnaires, 11interviews + transcripts	SPSS+WAV + TXT	8 GB	
Health (Nutrition)	No details available			

Table 4: Datasets to be deposited in UHRA

2.5 Improved compliance with funder requirements

The project team members have been actively involved in influencing developing of the policies and procedures in support of Research Data Management. These include involvement in:

- Developing the EPSRC Roadmap which has been confirmed at the UH Research Committee meeting on 31 May 2012 and submitted to the EPSRC .
- Revision of the UH data management policy and UPR IM12 – Appendix III for Research Data Management. As per new policy A Data Management Plan must be completed for every research project as follows:
 - a) if a Data Management Plan is required by the funding body, the Data Management Plan must be completed at the bid application stage;
 - b) If a Data Management Plan is not required as part of the bid application, then a Data Management Plan must be completed once the award is made and before the start of the research project.
- Contribution to the UH institution wide deliberations of a working group on open access, through reviewing standard licences including Creative Commons, Open Data Commons, Open Government License, UK Data Archive licence and other, and proposing Creative Commons Attribution-NonCommercial-ShareAlike 2.0 UK: England & Wales (CC BY NC SA) for our open data and Australian Governments Open Access and Licensing Framework (AusGOAL) for restrictive access.

Annex I

CASE STUDY 1: RDM for Professional Learning

Dr. Roger Levy of the School of Education, in the SSAHRI Research Institute, talks about data management for his research on the role of professional learning for early career education academics.

Research Area

Roger has been conducting research with UH staff on their experience of their first and second year of employment as academics coming from non-university contracts: that is, they were previously expert practitioners.

Funding

The project does not have formal funding: Roger carries out the research in his UH work-loaded research time.

Research Data

The data are eight interviews which are digital audio file recordings of semi-structured interviews with the participants. In the interviews the participants discuss their experience as early career academics and review their responses to a questionnaire they had filled in about their experiences of mentoring. Also collected are digital photographs of physical images the participants developed. Transcripts have been made of the recordings using Dragon software: Roger repeats the participants' words so that they can be automatically transcribed by the software.

Storage and Sharing

Storage is on Roger's passworded laptop. The data are also backed up on Roger's personal U drive on the university's server after each piece of data is collected or manipulated.

The data are not shared with anyone as they are of a potentially sensitive nature. The participants have been told, as part of the ethics protocol, that their data will not be disclosed to anyone else. The data would be very difficult to anonymise due to the small sample and identifiable origins.

Preservation

No decision has yet been made on the end date for data retention. Roger is still working on the research project so is still making use of the data. The data will be stored on the university servers for the foreseeable future.

Critical Issues

Throughout the project, the following issues and solutions were identified.

RDM Issue	Solution
Data is sensitive and the participant would be easily identifiable so sharing would be difficult	Keep secure and safe.
Data retention schedule: difficult to know how long it will be needed for.	Aim to keep it secure and safe on the university server for the future.
Data is in audio file format and in digital image format	Ensure that the format for storage is durable

CASE STUDY 2: RDM for Oilseed Rape Crops Research

Dr. Eren Demir talks about the UH project, led by Professor Bruce Fitt, in the School of Life and Medical Sciences, on predicting the global spread of phoma stem canker in oilseed rape crops . Eren is a member of the Hertfordshire Business School's Applied Statistics Unit, within the SSAHRI Research Institute. His role, along with other members of the team, is to carry out the statistical modelling for the project.

Research Area

The research aims to predict the global spread of phoma stem canker in oilseed rape crops. The study assesses both the potential risk and predicted economic losses of *Leptosphaeria maculans* being introduced to China. A paper, detailing the ground-breaking findings, has recently been published in Plant Pathology.

Funding

Twelve university-based partners are involved, including partners in China, Poland and the UK. The UH arm of the project is funded by the University of Hertfordshire, with £4,000 allocated to the Applied Statistics Unit for Eren and the team to carry out their role. Other partners self-fund their roles.

Research Data

The research data are in Excel format and include two datasets, one supplied by the Polish collaborators and one from Canada. Additional data have been collected in China. The data comprises observations of diseased and not diseased oilseed rape plants along with various explanatory variables for prediction. The data includes 100 observational records and uses statistical modelling techniques.

Eren and his team have developed models using disease spreading statistics with predictions relating to distance for the next 15 years.

Storage and Sharing

The raw data are kept on the UH networked S drive. The transformed data are kept on the Applied Statistics Unit team members' UH passworded computers. They are spread between five researchers who are each working on different files or parts of the programme.

Preservation

The data will stay on the S drive for the foreseeable future. The data are in Excel and SPSS formats and also backed up in ASCII format. This is suitable for a small dataset. No plans made yet for long-term preservation.

Critical Issues

Throughout the project, the following issues and solutions were identified.

RDM Issue	Solution
Need to share the data between members of a UH team while it is being worked on.	Use of UH shared networked S drive.
Need to share the data between partners in countries across the world, including China. Use of email currently.	Could use the DMS to share online.
If dataset gets larger may need alternative formats.	Eg MySQL.
Long-term preservation of the data – no plans	Store in the UH data archive
Sharing beyond the research team	Via the UK Data Archive.
Need to share drafts of papers detailing outputs in the lead up to journal submission. Currently use email.	Could use the UH DMS.

CASE STUDY 3: RDM for the Heritage Hub's HistoryPIN

Dr. Sarah Lloyd of the School of Humanities, University of Hertfordshire, in the SSAHRI Research Institute, talks about data management for her research on the building of links between the University of Hertfordshire and local history community groups in the region.

Research Area

Sarah's current research aims to identify and facilitate collaborations between community groups concerned with local history. Sarah's aim is to highlight relationships, to link the groups together and to enable UH's role in sustaining the groups' inter-relationships.

Funding

The Partners in History: collaborations in regional heritage project was awarded £20,000 by the AHRC. The HistoryPIN project will be using £6,500 of a £50,000 follow-up award for further Partners in History work.

Research Data

The data for this project will be entered directly by the community groups onto the repository, which is to be built by the independent, not-for-profit company, wearewhatwedo. Markers will be placed on a Google map with data which may include images, oral history recordings or pieces of information.

Storage and Sharing

The data will be shared on the HistoryPIN website <http://www.historypin.com/> in a tailored area to be designed specifically for Heritage Hub members. Only members will be able to post data onto the tailored area of the website but the contents will be on open access to the general public. The Heritage Hub and the HistoryPIN are sharing and collaborative projects, rather than extensive research data gathering projects, although academics may like to use the resource for sources of historical data.

HistoryPIN say that the institution or the individual owns the copyright of the data posted there. The owner needs to state the level of sharing permitted. The images sit on the HistoryPIN server but any videos or audio clips need to be uploaded to YouTube and a link included in HistoryPIN. Keywords should be used to aid searching for the

resources. HistoryPIN do not actively police the site but will react to complaints about content.

A smartphone app is also part of the HistoryPIN suite, which enables users to see the data as they walk around in a particular location. There is the potential for further dissemination of the data via social media such as twitter or facebook.

Preservation

Preservation depends on the HistoryPIN team maintaining the website. HistoryPIN is run by the not-for-profit company, wearewhatwedo, based in London, San Francisco and Bulgaria, which aims to use social tools and Google Maps to open up content to the public.

Critical Issues

Throughout the project, the following issues and solutions were identified.

RDM Issue	Solution
Preservation depends on the ongoing maintenance of the HistoryPIN website by the NFP company.	Back up data elsewhere too by asking contributors to send a copy for deposit on the DMS?
Potential for derogatory comments on the UH labelled tailored area	HH members to check regularly and report issues to the HistoryPIN administrators.
Volume and quality of data could be low without encouragement	UH team members to check regularly and encourage deposits by emails and meetings
Skills level of devolved depositors could be a barrier for posting media	Hold training and sharing of knowledge sessions
Rights issues need clarification. Potential for rights infringement by local groups when posting media – could UH be liable?	Check whether the institution, the individual posting the data, the NFP, or the originator of the medium owns the rights. Make sure that the HH partners are aware of the need to check rights.

CASE STUDY 4: RDM for 18th Century Business and Finance

Dr. Anne Murphy of the School of Humanities, in the SSAHRI Research Institute, talks about research data management for her research, which explores the connections between trust and state debt in the 18th century.

Research Area

Anne's current project focuses particularly on the use of the image of Britannia, as a strong symbol of trust, on coins, notes, ledgers, paperwork and so on. Her research questions ask why the image of Britannia? What does Britannia symbolise – what significance does her female gender have? How was Britannia used to represent public credit?

Anne also collaborates with a research partner, Dr Jennifer Basford, based at Queen Mary University, London who is a brand historian.

Funding

Anne was awarded £4,650 by the Economic History Society to carry out the research over a period of 7 months.

Research Data

The research data comprises a collection of instances of the image of Britannia appearing. These are mostly in the form of digital photographs in jpeg format of notes from ledgers and physical artefacts and images found in the British Library and the Bank of England Museum.

Anne's notes, made as Word documents and Excel spreadsheets, which relate to the artefacts and images are additional data.

Storage and Sharing

The data are stored on two laptops and backed up to an external hard drive and a memory stick. Anne also emails the data to herself as an additional back-up. There are no sensitivity issues as the data is publicly available historical data. Back up is not made to UH network drives due the difficulties of accessing this while working at home, as well as the need to share data with a research colleague at Queen Mary University.

The research outputs will include an article in a peer-reviewed journal and findings will also be made available to students via Anne's teaching and to other academics in the department for teaching purposes.

Preservation

The data will be kept by Anne for the lifetime of her career and may be further analysed by Anne for future projects. There is no strategy yet for preservation, other than in a computer folder and in hard copy.

Critical Issues

Throughout the project, the following issues and solutions were identified.

RDM Issue	Solution
Back-up of image and notes files	Use of several back-up locations but no networked storage. Consider making regular use of a UH networked drive, which could be shared with external research partners, for back-ups. Consider making use of passworded encryption for the memory stick.
Format of files	Consider storing the notes in a durable Open Data Format as back-up. Ensure that the format of the photographs is a reliable and durable format.
Metadata	Use metadata to label the images for future retrieval.
Preservation	Consider the UH Data Archive for long-term storage of the images and notes: rights issues would need to be investigated. Probably not a coherent enough dataset for the UK Data Archive as detailed explanation is needed.
Sharing of academic papers during collaborative write-up currently by email	Use of the UH Document Management system for storing notes and on-going academic papers with provision for inviting external collaborators.

CASE STUDY 5

RDM for Community Festival Sustainability

Dr. Allan Jepson of the Hertfordshire Business School, in the SSAHRI Research Institute, talks about data management for his research, which models community engagement and sustainability in community festivals and events.

Research Area

Allan's current research applies and critiques the Motivation-Opportunity-Ability Model to local community festivals with the aim of revealing factors that influence inclusive engagement within local community festivals. This RDM case study looks at research data management for Allan's recent study, carried out in partnership with Dr. Gill Ragsdell from the University of Loughborough and Professor Alan Clarke from the University of Pannonia. For this study, data were collected at the UtcaZene Festival, held in Veszprém, Hungary in 2012.

Funding

The Hertfordshire Business School awarded £1,500 In research Seedcorn Funding to Allan in 2012 to enable him to travel to Hungary to collect data. Planning, analysis and writing up has been carried out in the course of Allan's research allowance days at the HBS.

Research Data

Primary data was collected for this project in Hungary at the UtcaZene Festival, Veszprém. 175 anonymous responses were collected, from members of the festival audience, to paper-based questionnaires. These were administered at the festival by two specially trained Masters Students from the University of Pannonia. The questionnaires had been designed by Allan, ratified by the other two partners, pilot tested for usability by Gill at Loughborough, then translated into Hungarian by students at the University of Pannonia.

Three semi-structured anonymised interviews with key festival stakeholders were carried out in the Hungarian language by Dr. Aggie Rafé, an academic from the University of Pannonia, then translated by Aggie into English.

MP4 format videos and jpeg-format photographs of the festival were collected as additional data.

Storage and Sharing

In the field, the interview data were initially stored on a passworded iPad, and the images and videos on a passworded iPhone. Back at the University of Pannonia these were uploaded to a passworded computer and encrypted by 128-bit encryption during the upload.

The hard copy questionnaires collected by the students at the festival were put into special bags during collection, collected from the students at the end of the day by Allan and then locked into a cupboard at the University. The two Masters students, who had both signed a confidentiality agreement, inputted the data into Excel spreadsheets on Alan Clarke's computer.

The raw questionnaire Excel data files were emailed to Allan Jepson in an encrypted zip file, which was accessed once received using a separate password key file.

Allan continued to work on the data on his own UH passworded computer, being backed up to a passworded removable hard drive kept at his home. Students at the University of Loughborough also carried out SPSS analyses on the questionnaire data.

Preservation

No decision on long-term preservation has yet been made. The data are in current use for writing up academic journal and conference papers and will be used for subsequent comparison studies. An ESRC bid is being prepared to finance the data collection and analysis for the comparison studies.

Critical Issues

Throughout the project, the following issues and solutions were identified.

RDM Issue	Solution
Protection of data in the field	Hard copy stored in special bags; data files stored on passworded mobile devices kept on the person of the research partners.

Secure transfer of data between Hungary and 2 sites in England	Passworded encrypted zip file used.
Student research assistants used: potential for data misuse	Students signed confidentiality contracts and had been identified as trustworthy and reliable.
Dual language data: need to ensure that the content matches	Native-language Hungarian academic with fluent English collaborated on the project to translate the interviews and check the student translations
File formats	Files are currently in Excel, SPSS, Word, jpeg, and MP4. Although these are likely to be fairly long-lasting formats, the files should also be backed up in more durable formats such as OpenDocument Format (ODF), ASCII, tab-delimited format, comma-separated values, XML.
Back-up of data	Currently backed up at UH onto a hard drive. Recommend also backing up to Allan's networked U drive. Need to ensure that all partners and assistants are backing up to secure locations.
Synchronisation of data	Need to make sure that all partners are using the latest version of the data: check file naming conventions and metadata to ensure correct labelling.
Preservation of data	<p>A long-term preservation strategy needs to be decided upon. Data is currently being actively used so could not yet be released on open access. Further data are likely to be collected for comparison studies so the data will be in active use for several years. Consistent file naming policies will be even more vital as further data are collected.</p> <p>Future possibilities for long-term storage and for offering for open access are the UK Data Archive http://data-archive.ac.uk/</p>

	and the UH Data Archive. Sharing levels would need to be set, such as Creative Commons licences.
--	--

CASE STUDY 6: RDM for the Elstree Project

Howard Berry is Programme Leader in BA Film and Television, Senior Lecturer in Post-Production in the School of Creative Arts, which is part of the SSAHRI Research Institute. Howard talks about data management for his on-going oral history project, The Elstree Project.

Research Area

Howard's research explores the Elstree studio's filmmaking heritage. Howard is working in partnership with Bob Redman and Paul Welsh MBE of volunteer group Elstree Screen Heritage (ESH). The data comprises recorded stories collected from the props painters, the cameramen, the stills photographers, and the riggers who have worked at the studio during various periods since the 1940s.

Funding

The Elstree Project has no formal funding yet. The work is conducted during Howard's UH research time and the personal time of the Elstree Screen Heritage volunteers. Film students are using the project data collection to gain filmmaking experience for their course. Howard is currently planning a £60,000 AHRC bid for Contemo Portal media asset management software. A £15,000 bid has been made to SSAHRI for equipment to support the project. A PhD funding bid to enable Howard to draw on the data for a thesis is also in the planning stages.

Research Data

The data for this project currently comprise 33 interviews recorded in HD video format. Sixteen of the interviews have been transcribed so far: the transcription is carried out by the ESH volunteers. The audio and video elements of the interviews are recorded separately.

Howard has started tagging the video with metadata using QuickTime software but needs more complex professional-level software to complete this task more effectively. Individual frames or blocks of data need to be tagged. A set of standard metadata terms is being built.

Storage and Sharing

The video footage was captured on the original camera card which has been retained in unadulterated form. A duplicate has been made on computer in Quick Time format: this is an enduring format which is compatible with any computer system.

The data are legally shared between project partners UH and the Elstree Screen Heritage Volunteer Group under a signed agreement. Both parties would therefore need to agree where it should be deposited. Interviewees have signed an industry-standard release form.

The Project has recently been allocated a specific computer by the School of Creative Arts as one storage site. The data are also backed up on the UH shared servers and again within the School by Howard. Recordings are shared between UH and the ESH Chair using Dropbox as a brief transferal medium: the files are not left there for any length of time. Email transfer is not possible due to the size of the files.

The data have been shared with a UH PhD student and with a lecturer at Oxford University, who has used the content for her book. Data extracts are currently posted on Vimeo and one of the films has gained 95,000 hits after being promoted by Howard on Twitter and then picked up by a well-known film blogger.

Preservation

Preservation possibilities are to put the data on the UK Data Archive, to add the data to the Google Cultural Archive (they host content but the data provider retains copyright), or to put it on an iPad app with paid downloads. Howard would like the data to be as widely accessible as possible. Potential audiences are academics, film buffs, the general public and students.

Critical Issues

Throughout the project, the following issues and solutions were identified.

RDM Issue	Solution
Funding for effective data management	Apply for external funding UH to consider whether to make software available

Joint ownership of IPR	Ensure that legal contracts are discussed and fit for purpose from the beginning
Commercial value therefore exposure to 'theft'	Film online is watermarked and low quality so would not be suitable for commercial use

CASE STUDY 7: RDM for Early Language Intercultural Acquisition

Dr. Christina Schelletter of the School of Humanities, University of Hertfordshire, in the SSAHRI Research Institute, talks about data management for her research on the effects of learning another language at pre-school on first language acquisition.

Research Area

Christina's research is in the area of bilingual language acquisition and processing.

Funding

Christina is involved in a £9,480 European Commission Lifelong Learning funded research project in collaboration with partners from across Europe. A further bid for follow-up research has been submitted.

Research Data

The research data include the results of tests given to pre-school children. These are vocal tests, such as storytelling. Some are standard tests, others are project-designed tests. UH staff analyse the results of the tests, comparing them to norms where these exist. Test data relating to 200 children have been collected, together with parental data on first language, socio-economic class, education-level and whether the parents chose bilingual education.

Partners within the UK and across Europe have collected data in their local pre-school institutions. The European partners are based in Germany, Belgium and Sweden. UH has collected data from English mono-lingual children and from children in a bi-lingual school.

Statistical data are held in Excel and SPSS formats. Digital data are MP3 files. Transcripts are in Word format.

Storage and Sharing

Christina stores the data on her university computer, an external hard drive and her UH personal networked drive. Dropbox is used to exchange files between the partners and was used to store and work on the bid documents. Files are often too big to email.

The data may be shared in the future via the CHILDES (Child Language Data Exchange System) data base <http://childes.psy.cmu.edu/> This is part of the Talk Bank system for

sharing and studying conversational interactions and is free open access. Data posted on here is anonymised.

Preservation

Long-term preservation of the data is not currently written into the bid. The data are held at various institutions across Europe, making project-based preservation more difficult. The PI of the project is based in Germany.

Critical Issues

Throughout the project, the following issues and solutions were identified.

RDM Issue	Solution
Sharing of large files difficult, currently achieved via Dropbox	Use of UH Document Management System with access granted to other partners. Issue of UH not being the Principal Investigator, however.
Back-ups on an external hard drive	Ensure that the drive is encrypted and passworded.
Original format of data in Excel	Transform to Ascii or CSV for more durable format.
No preservation plan.	Create a Data Management Plan for the completed and future project.
Data shared with partners across Europe	Check partners' storage practices; store UH version of the data in the UH data archive. Store current data in the UH DMS.
Data need to be easily accessible by the partners in a granular format and in its raw format	Store in a variety of formats in the UH DMS.