Addressing data management training needs: a practice-based approach from the UK

Laura Molloy
Humanities Advanced Technology and
Information Institute (HATII)
University of Glasgow
Glasgow G12 8QJ
+44 (0)141 330 7133

Laura.Molloy@glasgow.ac.uk

Simon Hodson
Joint Information Systems
Committee (JISC)
5 Lancaster Place
London WC2E 7EN
+44 (0) 203 006 6071

S.Hodson@jisc.ac.uk

Stéphane Goldstein
Research Information Network (RIN)
20-24 Tavistock Sq
London WC1H 9HF
+44 (0)20 3397 3647

Stephane.Goldstein@researchi
nfonet.org

Joy Davidson
Humanities Advanced Technology and
Information Institute (HATII)
University of Glasgow
Glasgow G12 8QJ
+44 (0)141 330 8592

Joy.Davidson@glasgow.ac.uk

ABSTRACT

In this paper, we describe the current challenges to the effective management and preservation of research data in UK universities, and the response provided by the JISC Managing Research Data programme.

This paper will discuss, *inter alia*, the findings and conclusions from data management training projects of the first iteration of the programme and how they informed the design of the second, paying particular attention to initiatives to develop and embed training materials.

Keywords

Research data management; training; skills; digital curation; digital preservation; universities; research infrastructure; research support staff; postgraduate student research training.

1. INTRODUCTION: THE RESEARCH DATA MANAGEMENT CHALLENGE

The effective management of research data is an integral and inseparable part of the research process. Good research data management (RDM) therefore equates with sound research, a view which is reiterated by the Research Councils UK (RCUK) common principles on data policy [1] and the recent introduction by the UK's Engineering and Physical Sciences Research Council (EPSRC)'s explicit expectations for the management of research data generated by funded projects. Many other initiatives echo this view, such as the Organisation for Economic Co-operation and Development (OECD) principles for access to research data from publicly-funded research [2] and the ongoing efforts by Neelie Kroes, Vice-President of the European Commission responsible for the Digital Agenda for Europe[1] including the All European Academies (ALLEA) declaration on open science of April 2012 [3].

For the purposes of our discussion here, we are using the term 'data management' broadly to incorporate the notions of digital curation[2] and digital preservation[3], both as applied to research data produced by universities and other research institutions.

The challenge of achieving better RDM does not simply rely on addressing technical issues. These are important but tractable; equally important are organisational, policy and attitudinal issues. Universities are developing policies and technical infrastructure, but ultimately, researchers themselves have to be aware of the need for research data management, recognise that they have a role in managing their data, be willing to engage in RDM practice and have the skills, incentives and support to do so. Changes to the way research data is managed imply cultural change in the way research is practiced, whilst also continuing to support robust research processes. Disciplines vary in their levels of existing

---

[1]Described at http://ec.europa.eu/information_society/digital-agenda/index_en.htm

[2] See the Digital Curation Centre's definition at http://www.dcc.ac.uk/digital-curation/what-digital-curation

[3] See the Digital Preservation Coalition's definition at http://www.dpconline.org/advice/preservationhandbook/introduction/definitions-and-concepts

awareness and in common practice in the management and sharing of research data, so that researchers' behaviour is strongly influenced by their immediate environment as well as their disciplinary culture.

Researcher behaviour is specifically influenced by funders' requirements and increased recognition of the benefits of data sharing and re-use. Increased awareness in the researcher population of the options and advantages of data management and sharing can enable researchers to participate more fully in the emerging digital research economy.

The emerging digital research economy, as examined by the Royal Society's 'Science as a Public Enterprise' initiative, culminating in the 'Science as an Open Enterprise' report [4], has the potential to lead to significant advances in research and improve its overall quality by the provision of easier verification or reproducibility of the data underlying research publications. This can in turn allow new research questions to be asked of existing data, or integration of multiple datasets to achieve wider or more robust research conclusions. Research funders are naturally keen to obtain the greatest possible return on investment for grants disbursed. Increasingly, this is accepted to mean ensuring that research data is available for reuse and repurposing. Where the data is the product of unrepeatable observations, the case is easy to make. A well-known example is that there have been more papers published based on the reuse of archived data from the Hubble Space Telescope than those based on the use originally described when specific observations were requested.[4] Funders are increasingly aware that the potential of reuse can be extended to other research areas: as the research data management principles developed by the EPSRC state, sharing and promoting the reuse of research data is an important contributor to the impact of publicly funded research [5].

Journal editors are also sensitive to the need for research data to be available for verification and reuse. A growing number of journals are adopting increasingly stringent data availability policies. Most innovative and significant among these, perhaps, is the Joint Data Archiving Policy which underpins the Dryad Data Archive initiative [6]. Since August 2011, twenty BioMed Central titles have adopted data availability policies of varying rigour. Enthusiasm is growing around the idea of data papers and of data publications – or at the very least more effective linking and visualization of data through traditional publications. As the Opportunities for Data Exchange (ODE) Report on Linking Data and Publications [7] testifies, there are a growing number of innovative initiatives and in the next few years the publication of research data is likely to be recognized as a necessary part of the

publication of research results.[5] This has implications for the way in which researchers are trained.

Currently, a lot of publicly-funded data generated by UK universities is lost or inaccessible: this can cause serious difficulties for the researcher in the future when trying to re-access their data, and also greatly limits the possible return on the initial investment in that research. More sophisticated management of research data and improved linkage between research data and published outputs, then, clearly allows original research activity to be further exploited, yielding richer knowledge and wider impact. At the institutional level, there is currently a realisation in many universities that a significant change has to take place if these risks are to be controlled and these benefits are to be achieved.

2. RESPONDING TO THE CHALLENGES

In response to this set of challenges and the emerging new landscape within UK research, the JISC established the Managing Research Data (MRD) programme which has run as an ongoing set of activities since 2009. Through two iterations, the basic structure of the programme remains the same. Projects have tackled 'hard' (i.e. technical) and 'soft' (i.e. human) infrastructure challenges from the point of view of particular disciplines and, increasingly, specific institutions.

The work of the programme has addressed the practical and technical issues associated with research data management infrastructure and the challenges of data management planning, data citation, description, linking and publication. There has also been attention paid to the importance of training requirements. The first iteration of the programme (2009-11) funded five projects to address the training aspect which were supplemented by an additional support and synthesis project. The second iteration of the programme, launched in 2011, has a similar approach, again with a set of training-focused projects supported by an additional synthesis effort.

The programme included this training strand in both iterations in order to deliver a holistic approach to improving research data management in the university context in which both human and technical infrastructures are addressed.

3. THE JISC MANAGING RESEARCH DATA PROGRAMME

3.1 The Training Strand: Aims and Vision

It is a principle of the JISC MRD programme that there is little benefit in building systems and technical infrastructure unless motivation, recognition and reward and data management skills among the research population are also addressed. For this reason it was felt necessary for projects to develop and embed RDM training materials in discipline-focused postgraduate courses to help make clear the benefits and rewards of effective research data management at an early stage in the research career.

3.2 UK Researcher RDM Skills Needs

The shortfall in data management training in UK higher education is widely recognised. A 2009 *Nature* editorial 'Data's shameful neglect' concluded that 'data management should be woven into

---

[4] See http://archive.stsci.edu/hst/bibliography/pubstat.html. Observations by the Hubble Space Telescope are made on the basis of proposals, data is collected and made available to the proposers; data is stored at the Space Telescope Science Institute and made available after an embargo. Each year approx 200 proposals are selected from a field of 1,000; leading to c. 20,000 individual observations. There are now more research papers published on the bases of 'reuse' of the archived data than those based on the use described in the original proposal.

[5] This is another area of significant activity in the JISC Managing Research Data Programme, but one which goes beyond the scope of the present paper.

every course, in science, as one of the foundations of knowledge' [8], a view which has found agreement elsewhere [9].

This acknowledged need to increase skills in managing research data among staff in HEIs, including researchers, librarians and research support staff, was explored by the UKOLN 'Dealing with Data' report of 2007 [10] and Swan and Brown's 2008 report on 'The skills, role and career structure of data scientists and curators'[11]. It was considered further in the second Research Data Management Forum of November 2008. These discussions were presented in the form of a white paper by Graham Pryor and Martin Donnelly, where the case is forcefully made that 'data skills should be made a core academic competency' and that 'data handling [should be] embedded in the curriculum'.[9]

Some UK organisations have attempted to address this shortfall. The Digital Curation Centre (DCC) has developed a wealth of digital curation and research data management training materials.[6] The UK Data Archive provides extensive guidance and training materials on the creation, management and sharing of research data.[7] Under its Researcher Development Initiative, the Economic and Social Research Council (ESRC) funded a 'Data Management and Sharing for Researchers Training Programme' which developed a programme of training for researchers and research support staff [12].

Additionally, under the heading 'Information Management', the Vitae Researcher Development Framework (Vitae RDF) includes the following description of necessary skills acquisition: 'Develops a sustained awareness of the creation, organisation, validation, sharing and curation of data.' [13] An 'Information Literacy Lens' [14] on the Vitae RDF, which includes considerable emphasis on data management skills, has been developed in consultation with the Research Information Network (RIN)'s Information Handling Working Group.[8]

Research presented in the RIN-funded report *To Share or Not to Share* [15] highlighted researchers' concerns and misgivings about making research data available for verification and reuse. Early findings from projects in the JISC Managing Research Data programme, moreover, highlighted awareness and skills gaps among researchers and called for advocacy, guidance and training materials to address these issues.[9] Numerous reports have underlined the value of early intervention in the research career, including work by Sheila Corrall[10], the JISC and others[11].

---

[6] See http://www.dcc.ac.uk/training

[7] See http://www.data-archive.ac.uk/create-manage

[8] http://www.rin.ac.uk/our-work/researcher-development-and-skills/information-handling-training-researchers/working-group-i and see, e.g. http://www.vitae.ac.uk/researchers/1271-414711/Learn-about-information-handling-lens-on-Researcher-Development-Framework.html

[9] See project outputs for JISC Incremental: http://www.lib.cam.ac.uk/preservation/incremental/index.html, JISC Sudamih: http://sudamih.oucs.ox.ac.uk/, JISC MaDAM: http://www.library.manchester.ac.uk/aboutus/projects/madam/ and the University of Southampton: http://www.southamptondata.org/.

[10] Sheila Corrall has recognised the importance of data literacy training at postgraduate student level in 'Roles and

Consonant with such initiatives and the concerns they reflect, it has been observed that there 'is a need to go beyond the workshop and the short training course, and embed preparation for a professional (and personal) lifetime of digital data curation within the academic curriculum.'[12]

3.3 Research Support Staff RDM Skills Needs

As well as integrating research data management skills in curricula for discipline specialists, it is also necessary to develop targeted course materials for librarians, research support staff and data managers. Calls for the 'upskilling' of subject or liaison librarians for roles which encompass support for the management and preservation of digital research data have become more urgent of recent years. In 2008, Alma Swan and Sheridan Brown observed that 'The role of the library in data-intensive research is important and a strategic repositioning of the library with respect to research support is now appropriate'. Swan and Brown envisaged three roles for the library with regard to research data management as a precondition to data intensive research. These were:

1. Increasing data awareness among researchers.

2. Providing archiving and preservation services.

3. Developing a new professional strand of practice in the form of data librarianship.[11]

Such analyses of the field highlight the importance of addressing the respective needs of researchers, librarians and research support staff. The importance of training for librarians and research support staff was clearly recognized when designing the first MRD programme in 2009-10, but it was judged that other agencies and stakeholders were able to take forward work to develop training materials and curricula to improve research data management among librarians, for example. It was felt that the initial priority should be to address the needs of postgraduate students and early career researchers as relatively little work had been done in those areas. While this prioritization may have been reasonable, with the benefit of hindsight it is acknowledged that an opportunity was missed to advance work to improve data management skills among librarians and other key research support staff at that point. Work in the second iteration of the Managing Research Data Programme is designed to address this shortfall.

3.4 The RDMTrain Projects

In the first iteration of the JISC MRD programme, the object of the five training projects, collectively known as 'RDMTrain', was to create materials which translated, where possible, generic training resources into something meaningful and targeted to postgraduate students studying in specific disciplines, and viewed as an essential part of training and research skills in these

---

responsibilities: Libraries, librarians and data'. In G Pryor (Ed.), *Managing research data* (pp. 105-133). London: Facet.

[11] For examples, see the Arcadia project report of work in this area at University of Cambridge: http://arcadiaproject.lib.cam.ac.uk/docs/PINOTA-Report.pdf, and the recommendations of the JISC/RIN/DCC DaMSSI final report, available at http://www.rin.ac.uk/system/files/attachments/JISCfinalreport_DaMSSI_FINAL.pdf.

[12] Pryor and Donnelly 2009, p.166.

disciplines. These materials were to be sustained by embedding in existing postgraduate training provision as well as being made openly available through the Jorum portal[13].

The RDMTrain projects targeted the following disciplines: archaeology, the creative arts, geosciences, health sciences, psychology and social anthropology. A deliberate spread of disciplines was intentional: the programme did not intend to work only with scientific disciplines, which are often more familiar with discourse around the idea of data, but to also extend the terminology of data management into arts and humanities disciplines. The materials developed by the projects drew on user needs analysis undertaken with target audiences, and took the form of online guidance, practical software exercises, in-person training events and specimen data management plans alongside templates and supporting guidance materials.

3.5 The Data Management Plan

The RDMTrain projects of the first iteration of the MRD programme interrogated the DCC's Data Management Planning online tool [19] and its suitability for use within their target disciplines. They produced a set of discipline-focused templates for a data management plan (DMP), showing that discipline specificity, including the use of language appropriate to the discipline, encourages engagement with data management planning. However, further work is necessary to understand how data management planning can be optimised to the needs of a variety of disciplines and institutions.

The recent funding body mandates to embed data management planning as part of research practice can be useful to those providing training. Students wish to produce a DMP specifically relevant to them, often as a learning outcome of the course or as part of their wider skills development. Self-directed learning with access to customised guidance for the discipline and moderated exercises around the development of a DMP works well.

The DMP can be easily understood as another piece of administration which researchers are becoming obliged to complete. But the DMP can offer a research team a number of more sophisticated and engaging benefits when viewed as a dynamic tool which can be completed at the outset of the research work but regularly revisited during the work of the project to guide decision making about data use, re-use, storage and sharing. The DMP has potential as a pedagogical – or as one of the training projects suggested, andragogical – tool as, in order to be effective, data management planning must be an activity or learning process which draws on the experience of the working professional and informed by their experience in the role. Finding out the information required for the initial completion of the DMP helps the researcher to develop an awareness of the many issues connected to data management and leads to the ability for more sophisticated decision-making. This process can also provide a way to building the relationships between researchers and support staff which are required for the collaborative completion of the DMP; this can lead to new appreciation of the various roles involved in data management across the institution. In this way, the DMP also has the potential to influence researcher behaviour in regard to data management. In addition, the DMP is also a useful way of addressing the requirements of freedom of information legislation, by providing evidence of an intention to release research data[14].

The emphasis on data management planning is viewed by some funders and by the DCC as a core way of improving RDM practice. This seems a valid approach but there is still some work to be done on refining our understanding on what an optimal DMP – which aims to serve the requirements of a variety of stakeholders – might be.

3.6 DaMSSI Support and Synthesis

The five training projects of the first iteration were also accompanied by a support and synthesis project which was co-funded by the MRD programme and by the RIN, and was run with the co-operation of the DCC. This was the Data Management Skills Support Initiative ('DaMSSI') [16] which was overseen by the RIN Information Handling Working Group. One of DaMSSI's main purposes was to test the effectiveness of the Society of College, National and University Libraries (SCONUL)'s Seven Pillars of Information Literacy model [17] and the Vitae RDF for consistently describing data management skills and skills development paths in UK postgraduate courses. With the collaboration of the five projects, DaMSSI mapped individual course modules to the Seven Pillars and the Vitae RDF, and to the DCC digital curation lifecycle model [18] and identified basic generic data management skills alongside discipline-specific requirements. A synthesis of the training outputs of the projects was then carried out which investigated further the generic versus discipline-specific considerations and other successful approaches to training that had been identified as a result of the five projects' work.

In addition, DaMSSI produced a series of career profiles to help illustrate the fact that data management is an essential component - in obvious and less obvious ways - of a wide range of professions [16].

3.6.1 DaMSSI Findings and Recommendations

Finally, as a result of working with the RDMTrain projects, and in liaison with various wider stakeholders in data management and curation, DaMSSI formulated a set of recommendations for the institutions and projects embarking on future data management training development. These recommendations are based on synthesised feedback from the training strand projects about what factors contributed to the success of their training, and feedback received by the training projects from students whilst piloting their training offerings [19].

Some of the DaMSSI recommendations compared successful approaches in generic and discipline-specific approaches to data management training.

The first of these recommendations advised that those developing training work closely with disciplinary experts to ensure that terminology used within courses is accurate and clear to the target audience. This includes agreeing a basic definition of core concepts such as what 'data' can be within the discipline. This is particularly helpful for non-science disciplines.

---

[13] http://www.jorum.ac.uk

[14] See the guidance which specifies this requirement among others, from the Information Commissioner's Office, Sep 2011 - http://www.ico.gov.uk/news/latest_news/2011/ico-issues-advice-on-the-disclosure-of-research-information-26092011.aspx

Overviews and central descriptions of topic areas should be basic and generic, in order to introduce the topic at a level that is interesting but digestible for PhD students. This approach also allows modules to be more easily integrated into existing larger research methods courses.

In order to highlight relevance to the audience, however, generic material should be interlaced with discipline-specific examples, references and case studies wherever possible. This also helps to engage the audience, puts basic points into context and makes them understandable.

The RDMTrain projects found that training was more successful where training developers acknowledged accepted research practices within the discipline and worked to develop training materials that reflect these practices; for example, kinds of data handling, research funder expectations and popular archives and repositories.

Finally, training providers should use trainers with extensive knowledge of the discipline. Trainers who know the discipline well can provide the context and interlaced examples that engage students and make the topic seem relevant to them.

These observations raise important questions about training in research data management. Where, indeed, does such training ideally sit in the offering of a higher education institution, how is it most effectively delivered and who should be responsible for it? As a core research skill, intimately related to the practice of particular disciplines and affected by the specificities of the data collected, is it not right to argue that RDM should be tightly integrated with the postgraduate (or even undergraduate) training of a given discipline? Here for example, we might allude to the training in excavation and recording practice received by archaeologists, the knowledge of survey design and statistical analysis necessary among social scientists and the requirements among chemists and other experimental scientists to maintain a lab notebook. Is not RDM simply a core part of good research practice, which, along with other skills, should be inculcated early in the core disciplinary training of research students?

However, another point of view might be that RDM is a generic skillset, applicable to all disciplines. If RDM is regarded as a branch of information literacy, might it not be more effective and efficient to offer training alongside other such skills that are often delivered centrally, by staff that are specialists in approaches to information management? Recent studies [20, 21] of information handling among postgraduate students seem to suggest that there is a genuine, if not to say urgent, need for specific training in information handling skills and this cannot reliably be left to discipline specialists.

These considerations are fundamental and not susceptible to immediate solutions, particularly as we are at an early stage of integrating RDM training in curricula. Many universities will have to dose any solution with a generous helping of pragmatism. The JISC RDMTrain projects, DaMSSI, the RIN-led coalition and other stakeholders believe it is vitally important to promote RDM training and to share practice around delivery as this develops.

Another key group of the DaMSSI recommendations address the issues around language used in researcher training for data management. As identified in earlier JISC MRD programme work[15], the language and terminology used in the presentation of guidance and of training can make a significant difference in the extent to which researchers see the material as relevant to their discipline and engage with support infrastructure to better manage their data. The DaMSSI project found that, 'echoing the findings of the earlier JISC MRD Incremental project, many researchers don't understand much of the specialist language from the information or preservation worlds' [19]. These issues continue to be explored in the work of the JISC-funded SHARD project.[16]

Language issues arose again when DaMSSI worked with the training projects to ascertain use cases for the SCONUL Seven Pillars and Vitae Researcher Development Framework models. In the first instance, many project staff members were confused by the acronym 'RDF' for the Researcher Development Framework, this acronym already being widely understood in this community to denote a completely different concept. In addition, each of the Seven Pillars has a name that has immediate relevance to data management, but the definition of these terms is at times different for different audiences. For example, the 'Plan' pillar in the Seven Pillars model focuses specifically on search strategies for locating information, whilst 'plan' within a data management lifecycle has a broader and earlier definition of planning how data will be managed at the same time as a research project is outlined. That process, however, would currently be more aligned within the Seven Pillars model with the 'Scope' pillar.

DaMSSI recommended that training providers should avoid using acronyms and data curation-specific terminology, and instead explain principles and issues in language that is understandable to a general audience and is not already weighted for the audience's discipline: for example, the term 'curation' already has specific meaning for much of the creative arts.

It is hoped that these recommendations will contribute to the subsequent development of successful postgraduate-level RDM training materials.

4. FUTURE ACTIVITY

Activities in the second JISC Managing Research Data Programme to address training requirements are driven by the findings of the first programme and the recommendations of the DaMSSI project. There has also been an effort to cover areas relatively neglected in the first programme and to respond to changing circumstances.

4.1 Future RDM Responsibilities: Cross-Campus

The EPSRC's 'Policy Framework' and 'Expectations' [23] for RDM have underlined what was already a growing recognition that solutions to the research data challenge will require 'cross-campus' responses which coordinate a number of stakeholders, including researchers, the library, computing services and research support services. Although much responsibility for research data management must necessarily remain with the individual researcher, PI or research group, it has been recognized that various agencies within universities and research institutions

---

[15] Namely the JISC Incremental project at the Universities of Cambridge and Glasgow, project website available at: http://www.lib.cam.ac.uk/preservation/incremental/index.html.

[16] This project blogs at http://shard-jisc.blogspot.co.uk/

have important supporting roles to play. This realisation coincides with increasingly urgent calls for university libraries to adapt to the requirements of research as it becomes more data-centric. The recent report by Mary Auckland for Research Libraries UK, appropriately entitled *Re-Skilling for Research*, communicates a palpable sense of urgency:

> A shift can be seen which takes Subject Librarians into a world beyond information discovery and management, collection development and information literacy training, to one in which they play a much greater part in the research process and in particular in the management, curation and preservation of research data, and in scholarly communication and the effective dissemination of research outputs. [24]

The Dutch 3TU Datacentrum, a collaborative effort between the Netherlands' three technical universities, has developed the 'Data Intelligence 4 Librarians' course for which there is a substantial amount of online material [25]. The programme aims to equip librarians better to 'to advise researchers effectively and efficiently' in data curation. Such work is extremely useful, but there remains – as in the case of researchers themselves – a need to embed training in research data management skills in Library and Information Science postgraduate courses in order to ensure such skills are a *sine qua non* for the next generation of librarians. With these issues in mind, the Managing Research Data programme has, in its second iteration, explicitly targeted the development of training materials for librarians, funding a project led by the University of Sheffield iSchool.

4.2 RDMTrain 02

By and large, the training projects in the first Managing Research Data programme focused on the arts, humanities and social sciences. This orientation stemmed from a number of related considerations: the opportunity to build on existing materials coincided with a tangible need for skills development and an estimation that the challenges in these subject areas, while significant, may yet be relatively tractable. There has also been feeling that the more technical focus of STEM subjects – and the higher levels of funding available – meant that JISC-funded work was less necessary and would have a less tangible impact. However, reports such as the RIN study into *Information Practices in the Physical Sciences* [26] suggest that such assumptions may, at least in part, be misplaced. The second iteration called for projects to develop materials in subject areas which had not been covered in the first programme, and it is notable that projects targeting more technical subjects were prominent among those funded and include computer science, digital music research, physics and astronomy.

4.3 DaMSSI-ABC

As a whole, and specifically through the new DaMSSI-ABC support project, the training strand of the JISC Managing Research Data programme seeks to promote the incorporation of RDM components into the training of current and future researchers and research support staff. Building on the findings and recommendations of the first programme, the second iteration seeks in particular to ensure that materials are as reusable as possible and to promote them with learned societies and professional bodies.

'ABC' in the support project's name stands for Assessment, Benchmarking and Classification, underlining a commitment to ensuring that the training materials developed are as discoverable, reusable and interoperable as possible. With the assistance of the support project, the programme will work closely with the JISC-funded Jorum repository for reusable learning and teaching materials (or in Jorum terminology, Open Educational Resources). In collaboration with the MRD programme, Jorum will be piloting a research data management-flavoured portal in order to assist access to training materials [27]. The motivation behind this activity is to a) draw attention to research data management as an important component of more general research skills, and b) make the related materials more easily discoverable and reusable.

An essential component of reusability, when it comes to learning and teaching resources including training materials, is understanding precisely how the material might be incorporated into existing courses of diverse characteristics. Standardised descriptions, mapping of assessment, benchmarking required attainments and detailing subsequent classification are arguably the necessary components for the interoperability of training materials. A central focus of the DaMSSI-ABC project will be to make practical and informative recommendations on the basis of examining UK and international frameworks for benchmarking and classifying training materials. Existing models include the US Library of Congress's Digital Preservation Outreach and Education (DPOE) audience classification pyramid[17] (which may provide a useful guide for identifying courses aimed at executive-level strategic planners, operational managers and practitioners) and the Vitae RDF, but other initiatives will be taken into account, as well as the expertise of key UK stakeholders.

4.3.1 DaMSSI-ABC: The Role of Learned Societies

The important role of learned societies and professional bodies in contributing to the formulation of training materials, endorsing them and promoting them as part of the development support that they offer to their members is clearly recognised. As custodians of professional standards, these bodies are obvious interlocutors for the purpose of helping to promote data management skills, and to get these skills better recognised by students and researchers as indispensable elements in their career development. However, most such bodies have had little or no involvement in information literacy issues. The DaMSSI-ABC project in its support role will work to encourage and facilitate a dialogue between the funded projects and appropriate learned societies / professional bodies. This work will aim to ensure that data management skills are recognized by relevant learned societies and professional bodies as an indispensable part of researchers' career development, to accordingly identify data management champions within these organizations and to involve them in identifying means of skills assessment and benchmarks.

4.3.2 DaMSSI-ABC: Other Planned Activity

Other principle areas of activity include:

- Encouraging the early encounter with research data management issues in the research career;

- Working to help researchers and research support staff to plan career development;

---

[17] A useful description and diagram of the DPOE pyramid, along with definitions of each audience, is available at http://www.digitalpreservation.gov/education/educationneeds.html

- Exploring ways to assess and compare courses; and,

- Reporting, where possible, on diverse strategies for incorporating RDM training in discipline-specific curricula or more generic research skills offerings.

In this way, the DaMSSI-ABC project aims to contribute to the uptake and reuse of RDM training materials in the UK (and potentially internationally) as well as increasing our understanding of the most effective description, benchmarking and classification of such materials.

5. CONCLUSION
This paper relates the efforts of the JISC Managing Research Data programme to encourage the development and uptake of RDM training materials across UK institutions and disciplines. Although much progress has been made, the authors are obliged to recognize that considerable work is still required before research data management training is widely incorporated into postgraduate training curricula.

It is hoped that this paper will contribute to an international debate around the place of research data management training, and how it may best be delivered and promoted. We particularly emphasise the value of emerging work a) to engage learned societies and professional bodies; b) to establish practical and effective means of describing, benchmarking and classifying training materials to promote reuse; and c) to encourage colleagues across the university campus to engage with research data management and to tackle its challenges in collaboration.

6. ACKNOWLEDGMENTS

7. REFERENCES
[1] Research Councils UK (2011). *RCUK Common Principles on Data Policy.* Available at: http://www.rcuk.ac.uk/research/Pages/DataPolicy.aspx

[2] Organisation for Economic Co-operation and Development (2007). *OECD Principles and Guidelines for Access to Research Data from Public Funding.* Available at: http://www.oecd.org/dataoecd/9/61/38500813.pdf

[3] European Federation of National Academies of Sciences and Humanities (2012), 'Joint declaration as European Commission VP Neelie Kroes and ALEA agree on Open Science, Rome, 11-12 April 2012'. Press release, published 11th April 2012. Available at: http://www.allea.org/Pages/ALL/33/144.bGFuZz1FTkc.html

[4] The Royal Society (2012). *Science as an Open Enterprise: Final Report.* Available at: http://royalsociety.org/policy/projects/science-public-enterprise/

[5] EPSRC (2011). *Policy Framework on Research Data: Principles.* Available at: http://www.epsrc.ac.uk/about/standards/researchdata/Pages/principles.aspx

[6] Dryad (2011, updated 2012). *Joint Data Archiving Policy.* Available at: http://datadryad.org/jdap

[7] Opportunities for Data Exchange (2011). *Report on Integration of Data and Publications.* Available at: http://www.alliancepermanentaccess.org/wp-content/uploads/downloads/2011/11/ODE-ReportOnIntegrationOfDataAndPublications-1_1.pdf.

[8] *Nature* (2009). 'Editorial: Data's shameful neglect'. 461, p. 145. Published online 9 September, 2009. Available at: http://www.nature.com/nature/journal/v461/n7261/full/461145a.html. doi:10.1038/461145a

[9] Pryor G and Donnelly M (2009). 'Skilling Up to Do Data: Whose Role, Whose Responsibility, Whose Career?' *International Journal of Digital Curation,* Vol. 4, No. 2, pp. 158-170. doi:10.2218/ijdc.v4i2.105

[10] Lyon L (2007). *Dealing with Data: Roles, Rights, Responsibilities and Relationships*, UKOLN consultancy report. Available at: http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing_with_data_report-final.doc

[11] Swan A and Brown S (2008). *The Skills, Role and Career Structure of Data Scientists and Curators: an Assessment of Current Practice and Future Needs: Report to the JISC.* Available at: http://www.jisc.ac.uk/media/documents/programmes/digitalrepositories/dataskillscareersfinalreport.pdf

[12] Lyon L, Van den Eynden V and Bishop E (2011). ESRC Researcher Development Initiative, 'Data Management and Sharing for Researchers Training Programme'. Outputs available at: http://www.esrc.ac.uk/my-esrc/grants/RES-046-25-0038/read

[13] Vitae (2011). *Researcher Development Framework.* Available at: http://www.vitae.ac.uk/policy-practice/165001/Researcher-development-framework-consultation.html

[14] Vitae (2012). *Information Literacy Lens on the Researcher Development Framework.* Available at: http://www.vitae.ac.uk/CMS/files/upload/Vitae_Information_Literacy_Lens_on_the_RDF_Apr_2012.pdf

[15] Swan A and Brown S (2008). *To Share or Not to Share: Publication and Quality Assurance of Research Data Outputs. A report commissioned by the Research Information Network (RIN).* Available at: http://rinarchive.jisc-collections.ac.uk/our-work/data-management-and-curation/share-or-not-share-research-data-outputs

[16] Research Information Network (2010). 'Data Management, Information Literacy and DaMSSI'. Available at: http://rinarchive.jisc-collections.ac.uk/our-work/researcher-development-and-skills/data-management-and-information-literacy

[17] SCONUL (2011). *The SCONUL Seven Pillars of Information Literacy Core Model for Higher Education.* Available at: http://www.sconul.ac.uk/groups/information_literacy/publications/coremodel.pdf

[18] Digital Curation Centre (2010). *The Digital Curation Lifecycle Model*. Available at: http://www.dcc.ac.uk/resources/curation-lifecycle-model

[19] Molloy L and Snow K (2011). *DaMSSI Final Report*. Available at: http://www.dcc.ac.uk/webfm_send/532 and http://rinarchive.jisc-collections.ac.uk/system/files/attachments/JISCfinalreport_DaMSSI_FINAL.pdf

[20] Research Information Network (2009). *Mind the Skills Gap: Information-Handling Training for Researchers.* Available at: http://rinarchive.jisc-collections.ac.uk/our-work/researcher-development-and-skills/mind-skills-gap-information-handling-training-researchers

[21] JISC and the British Library (2012). *Researchers of Tomorrow: The Research Behaviour of Generation Y Doctoral Students*. Available at: http://www.jisc.ac.uk/publications/reports/2012/researchers-of-tomorrow.aspx

[22] Digital Curation Centre (2007). *Data Management Online Planning Tool*. Available at: https://dmponline.dcc.ac.uk/

[23] Engineering and Physical Sciences Research Council (2011). *EPSRC Expectations*. Available at: http://www.epsrc.ac.uk/about/standards/researchdata/Pages/expectations.aspx

[24] Auckland, M (2012) for Research Libraries UK. *Re-skilling for Research.* Available at: http://www.rluk.ac.uk/content/re-skilling-research

[25] 3TU.Datacentrum (2012). *Data Intelligence 4 Librarians* course available at: http://dataintelligence.3tu.nl/en/news-events/news-item/artikel/data-intelligence-4-librarians-online/

[26] Research Information Network (2011). *Collaborative yet Independent: Information Practices in the Physical Sciences*, report for RIN and the Institute of Physics. Available at: http://www.rin.ac.uk/our-work/using-and-accessing-information-resources/physical-sciences-case-studies-use-and-discovery-

[27] Siminson N (2012). 'Piloting institutional and subject flavours of Jorum'. Blogpost, published online 19 April 2012. Available at: http://jorum.ac.uk/blog/post/30/piloting-institutional-and-subject-flavours-of-jorum