Chapter Eight

Fostering Open Science Practice through Recognising and Rewarding Research Data Management and Curation Skills

Joy Davidson
Digital Curation Centre, UK
Email: joy.davidson@glasgow.ac.uk

Review status: peer reviewed

How has the research landscape changed in the UK?

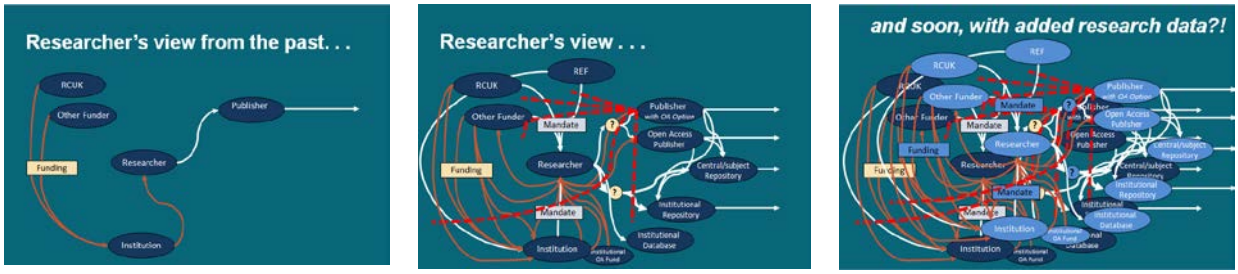
Not too long ago, researchers primarily needed to worry about securing grant income, doing their research, and publishing their findings in an appropriate, subject specific journal. The process was straight-forward and there was little need to report activity in detail at the institutional level. However, with the advent of the Research Assessment Exercise (RAE) back in 2008, things began to change as research publication metrics relating to both quantity and quality had to be produced by participating Higher Education Institutions (HEIs). The RAE was:

> conducted jointly by the Higher Education Funding Council for England (HEFCE), the Scottish Funding Council (SFC), the Higher Education Funding Council for Wales (HEFCW) and the Department for Employment and Learning, Northern Ireland (DEL). The primary purpose of the RAE 2008 was to produce quality profiles for each submission of research activity made by institutions (HEFCE, SFC, HEFCW & DEL. 2008).

In the run up to RAE 2008, many HEIs introduced institutional repositories and associated deposit policies to make the reporting process more accurate and efficient. A few years later, the emergence of Research Councils UK's *Common Principles on Research Data Policy* (RCUK, 2011) placed further demands upon researchers in terms of managing and sharing outputs to make research more accountable, open, and reusable. As part of this new landscape, data management plans (DMPs) are now required at the grant application stage by the majority of Research Councils UK (RCUK) funders. The only exception is the Engineering and Physical Sciences Research Council (EPSRC) who do not want to see a DMP alongside the grant application but assume that one will exist locally. Bill Hubbard, Director of the Centre for Research Communications at the University of Nottingham, has captured the increasing complexity of the changing research landscape in Figure 1.

The majority of RCUK funders place the onus on researchers to provide evidence that research data management and sharing are being considered from the outset of new projects via the completion of a data management plan (DMP). However, things changed when EPSRC released its Policy Framework on Research Data in 2011. The policy included nine expectations for those in receipt of - or seeking to be in receipt of - EPSRC funding. Crucially, EPSRC placed the onus on the research institution to demonstrate that suitable infrastructure was in place rather than on the individual researcher. The deadline for compliance with EPSRC's nine expectations came into effect on May 1, 2015 and, in early June of the same year, a light-touch survey was issued to senior management in UK HEIs to assess progress. EPSRC's policy has been absolutely instrumental in unlocking institutional funding for the development of fledgling RDM support services and systems and has had practical implications for both researchers and support staff. The follow up to the 2008 RAE - the Research Excellence Frame-

Figure 1: View of the changing research landscape for researchers (Hubbard, 2015)



work (REF) - has complicated the research landscape even further. Now HEIs are required to measure and report on the impact of their research activities as well as the quantity and quality of their research outputs. Impact accounted for 20 per cent of the overall REF score in 2014. In the guidance provided by HEFCE, SFC, HECW, and DEL, Research Excellence Framework (REF) review panels were asked to assess the:

> reach and significance of impacts on the economy, society and/or culture that were underpinned by excellent research conducted in the submitted unit [of assessment], as well as the submitted unit's approach to enabling impact from its research (HEFCE, SFC, HEFCW & DEL, 2014).

While the UK has been a leader in terms of driving forward the open science agenda through various governmental and funding body mandates, it is important to note that this movement is truly global in nature. Funding bodies in the US have also been early implementers of the open science movement and the European Union and South Africa are currently introducing mandates relating to the management and sharing of research data and publications.

To ensure that researchers and HEIs can operate effectively in this new research landscape, far greater cooperation and collaboration among the entire range of stakeholders involved in the research lifecycle is needed from the very outset of new research projects.

What support services will researchers need to be effective in this new research landscape?

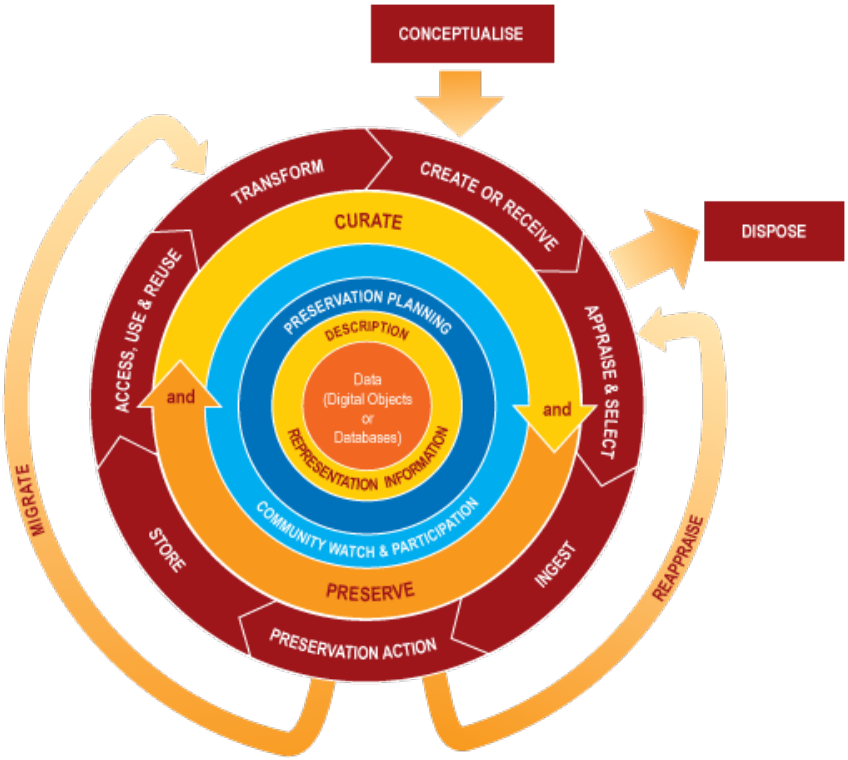
To equip researchers to embrace open science practices, HEIs need to ensure that adequate support systems and services are in place. Established in 2004 and funded by JISC, the Digital Curation Centre (DCC) helps UK universities build capacity and capability for research data management and curation. In 2007, the DCC developed its Lifecycle Model (Higgins, 2008) to assist in breaking down the range of roles, responsibilities and actions that need to be considered as part of good research practice.

In a practical sense, the model has helped to progress the development of procedures, tools, and support services for research data management and sharing within institutions. The Lifecycle Model has also been used to inform the development of digital curation curricula across the globe. However, while good progress has been made in defining services and educational programmes, there is still much to be done to embed the necessary research data management and curation skills into the daily workflows of all researchers and support staff.

Over the past ten years, the DCC has been involved in numerous international initiatives to help improve the quality and consistency of research data management and curation training and education. This has included participation in working groups to define digital curation curricula and leading several research projects to define, assess and benchmark skills required for RDM, curation and open science (for example: Research Data Management Skills Support Initiative (DaMSSI) in 2010; DaMSSI-Assessment, Benchmarking and Classification in 2011; Facilitate Open Science Training for European Research (FOSTER) in 2014).

A number of skillsets and aptitudes have been defined to reflect the range of competencies

Figure 2: Digital Curation Lifecycle Model (DCC Curation Lifecycle Model, 2015)



needed by different stakeholders across the curation lifecycle including ARMA's *Professional Development Framework for Research Administrators* (2011), CILIP's *Professional Knowledge and Skills Base for Information Professionals* (2013) and *Vitae's Information Literacy Lens on the Researcher Development Framework for Researchers* (2012). However, there has been a lack of standardised approaches that enable contributions to the research processes from a range of stakeholders to be represented consistently when reporting. Accordingly, it has been very difficult to see how – or if – curation skills acquired through formal education, professional development training, or on-the-job learning have translated into improved daily workflows across the research lifecycle. In addition, there are few reward structures in place that recognise those who are making use of their research data management and curation skills to undertake or to support open science. This gap in the recognition of curation skills can often make it difficult to incentivise researchers or support staff to put the necessary time and effort into acquiring these skills.

How the CRediT taxonomy could help to join the skills acquisition and skills recognition gap

In light of the fact that many publishers now require the roles of authors listed in academic papers to be clarified, the CRediT contributor roles taxonomy was developed by a small group of journal editors, Harvard University, and the Wellcome Trust in 2012 to help make contribution reporting more consistent (Allen et al, 2014). Project CRediT is:

> …led by The Wellcome Trust and Digital Science, facilitated by CASRAI and NISO and supported by the Science Europe Scientific Committee for the Life, Environmental and Geo Sciences. The project is developing and maintaining recommendations for a science-oriented contributor role taxonomy and its implementation (Project-CRediT, 2015).

The 14 term taxonomy that was developed aims to describe a more accurate and complete range of the contributions involved in producing research publications. The taxonomy was tested between August and November 2013 through an online survey of researchers who had recently

contributed towards academic publications. The taxonomy also helps to specify the degree of contribution (i.e., lead, equal, or supporting) more consistently. The taxonomy could help to provide evidence of research data management and curation skills being applied in a range of research and support roles and, by enabling these efforts to be recognised more formally by publishers and employers, help to incentivise and drive forward good practice.

In addition to being applied retrospectively to recognise contributions, there is scope to apply the taxonomy proactively when designing and describing educational programmes and professional development training courses in all disciplines. The proactive use of the taxonomy might lead to more effective course selection and the ability of participants - and their employers - to measure and reward learning outcomes over time. By employing a standard set of terms, the taxonomy may also enable better alignment between disciplines so that the skillsets of a range of stakeholders are more complementary. The alignment of curation related skillsets between disciplines has proven very difficult to date (Davidson et al, 2012).

By acknowledging and recognising that many different roles are involved in producing research outputs and employing a standard set of terms that describes these various contributions, the mapping of research data management and curation skills for researchers, librarians, administrators and other stakeholders is far more feasible. The potential value of the CRediT taxonomy for this purpose is currently being explored through the Research Data Alliance Interest Group on Education and Training (Davidson and Jung, 2015).

CRediT Taxonomy

The next section of the paper will introduce the 14 CRediT taxonomy elements as they relate to the Digital Curation Lifecycle Model and make specific recommendations relating to the potential roles of Library and Information Science (LIS) professionals. Where appropriate, multiple CRediT taxonomy elements have been clustered together under specific lifecycle stages.

Conceptualisation stage

Coming up with research questions, seeking suitable funding calls, and developing pre-award data management planning should all take place during the conceptualisation stage of the curation lifecycle. It is important to consider all of the relevant roles and responsibilities relating to data sharing to ensure that adequate funding can be requested for any data cleaning or anonymisation that might be required. Table 1 reflects the relevant CRediT taxonomy elements for the conceptualisation stage .

Research in all disciplines is becoming increasingly data driven. This means that during the conceptualisation stage of the research project, librarians may start to play an important role in helping researchers carry out a data review to progress the development of new research ideas. This draws upon the established role of librarians in supporting literature reviews, but will help to ensure that researchers have undertaken a landscape analysis to ensure that any proposed data collection activity will not be duplicating data that already exists. Evidence that new research projects will not be duplicating data collection is something that some funding bodies are now actively seeking in grant applications, as can be seen in this extract from the Economic and Social Research Council (ESRC) data management planning guidelines:

> Where research grant applicants plan to create new data as part of their ESRC-funded proposal, they must demonstrate that no suitable data are available for re-use. ESRC encourages the re-use of existing data and therefore encourages applicants and grant holders to consider the breadth of data available from various sources before committing to primary data collection (ESRC, 2015).

In cases where data do already exist, it might be the case that they are not suitable for reuse due to poor quality, lack of contextual metadata, or unsuitable data licences being applied. In-

Table 1: CRediT taxonomy elements for the conceptualisation stage (Allen et al, 2014)

#1 conceptualization	Ideas; formulation or evolution of overarching research goals and aims.
#8 data curation	Management activities to annotate (produce metadata), scrub data and maintain research data (including software code, where it is necessary for interpreting the data itself) for initial use and later reuse.
#12 supervision	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.
#13 project administration	Management and coordination responsibility for the research activity planning and execution.
#14 funding acquisition	Acquisition of the financial support for the project leading to this publication.

deed, in some cases, it may not even be possible to search for, or identify, existing research data held by HEIs. In such cases, librarians may be crucial in helping to document searches that have been carried out and in making the case in the grant application that re-creating data is necessary.

Librarians will also play a key role in contributing to the development of required data management plans for new grant applications and many library staff in UK HEIs are actively involved in developing customised, institutional guidance for inclusion in DMP tools (e.g. DMPonline). Effective curation starts with good data management planning and, in this respect, librarians will play a vital role in ensuring that valuable data are managed and curated over the entire digital curation lifecycle. Indeed, in the UK, many of the early data management and curation support services emerging in HEIs are being led by the library. In addition to general research data management planning support, librarians will be influential in determining current RDM and curation practice and capacity within the institution and in identifying future requirements using methodologies such as the Data Asset Framework (DAF).

Librarians' expertise around making information accessible and understandable will be crucial for longer term curation and reuse. In terms of improving the visibility of research data generated by the institution, librarians will be instrumental in defining requirements for, and the establishment of, institutional data catalogues that will be harvested by national data discovery services such as Research Data Australia and Jisc's Research Data Discovery Service.

Research data management and curation requires dedicated effort and the allocation of resources. Fortunately, as stated in the *RCUK Common Principles on Research Data Policy*, funders agree that the use of public funds to support data accessibility and reuse are appropriate. Clarifications on the sorts of costs that might be eligible in grant applications were provided by RCUK following the DCC's *Research Data Management Forum on Funding for Research Data Management* (Ryan, 2013). While a range of RDM and curation activity costs are eligible within grant applications, they must be clearly justified – just as any other budget request would need to be. There is a real risk though that these eligible costs may be missed unless effective data management planning is carried out early and in consultation with key stakeholders. Librarians have a proven track record as trusted liaison points between researchers and central university services. By continuing to foster this pivotal role, librarians could play a crucial role in ensuring that adequate RDM and curation activities are identified and costed into new grant applications. Tools such as the *UK*

Data Archive's Data Management Costing Tool and Checklist (UK Data Archive, 2013), the DCC's *Five Steps to Decide What Data to Keep* (DCC, 2014) and the *Collaboration to Clarify the Costs of Curation* (*4C*, 2013 and 2015) will be useful references for librarians to consult when costing curation activities.

In the longer term, librarians will play a key role in developing and delivering effective training for research staff and students around research data management and curation themes, for example the MANTRA research data management course which was developed by EDINA at the University of Edinburgh (2010).

Create or receive stage

During the creation stage of the curation lifecycle, the data management plan should be updated with more concrete information regarding specific research approaches and refer to the adoptions of relevant disciplinary standards. During the active stage of the project, researchers will be collecting and/or using data and performing analyses. In many cases, they will also develop code to enable them to perform specific analyses and/or to visualise their data more effectively. Maintaining these additional research outputs is essential for longer term reuse – both by those who created the data and by others. Relevant CRediT taxonomy elements for the creation stage can be seen in Table 2.

Librarians may have an active role to play in helping to define methodologies around data collection in new research projects. While researchers will be adept with disciplinary norms and standards of good practice within their particular community of practice, librarians will be influential in advising on broader metadata standards – particularly in light of making research outputs more visible, understandable and ultimately reusable to other disciplines as well as by the general public where appropriate. In a recent article, Markin highlights that:

> Many libraries are playing a leadership role in this effort. Data storage and preparation can get complicated quickly for the nonexpert, and it's useful to have someone knowledgeable who can guide you through the intricacies of naming and formatting data for easier access by others (Markin, 2015).

For researchers working with sensitive data, librarians may be well positioned to spot potential methodological conflicts between data protection and plans for data sharing and reuse as outlined in data management plans and/or pathway to impact statements. Such risks could then be flagged to researchers and ethics teams to ensure that suitable consent forms and terms and conditions for access and reuse are developed for the data.

In relation to this, librarians could play a role in helping researchers to assess the level of data cleaning and/or anonymisation that may be required to facilitate data sharing and reuse. Data cleaning and anonymisation can be costly activities and, as noted above, librarians may be pivotal in ensuring that additional costs relating to RDM and data sharing are identified and factored into grant applications. However, it is currently difficult to get recognition for this valuable work. The recent Expert Advisory Group on Data Access (EAGDA) report stated that:

> Understanding user requirements and extracting, formatting, annotating and cleaning data to maximise its utility and value to other users can take up a significant proportion of staff time. These efforts are often considered to be activities that should be undertaken as "part of the day job" yet are rarely recognised or credited academically (Expert Advisory Group on Data Access, 2015).

In cases where data cleaning or anonymization costs are prohibitive, librarians may be called upon to offer advice on alternative approaches that will meet funders' data sharing expectations while respecting data protection laws. For instance, if a proposed social science research project aims to conduct video-taped interviews as part of the data collection process and the researcher has indicated that they would like to

Table 2: CRediT taxonomy elements for the creation stage (Allen et al, 2014)

#2 methodology	Development or design of methodology; creation of models
#6 investigation	Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection
#3 software	Programming, software development; designing computer pro-grams; implementation of the computer code and supporting algorithms; testing of existing code components
#5 formal analysis	Application of statistical, mathematical, computational, or other formal techniques to analyse or synthesize study data.
#11 visualization	Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.

share an anonymised version of this data after the project ends, it may be the case that the costs of anonymising the video data would be greater than the actual award is worth. In such cases, it may be more appropriate to state in the data management plan that an anonymised transcript will be shared as an output of the project rather than the video data itself.

Researchers often develop bespoke software to enable them to carry out specific analyses and/or to create visualisations of the data they are working with. To ensure that research findings are reproducible, access is needed not only to publications and the underlying data, but also the software, algorithms, and techniques that were employed to perform these analyses to enable validation and effective reuse. In some cases, the software developed to analyse or enable a visualisation of the data is more at risk than the research data itself. This is particularly true for researchers who are developing new tools and techniques for analysing existing, well curated data sets (e.g., historical weather data held by National Oceanic and Atmospheric Administration (NOAA)). In the same way that data management planning helps to ensure that the management of research data is considered from the outset of new research activity, the development of a software sustainability plan helps to ensure that any required software or code is maintained throughout the project and beyond. Michael Jackson of the Software Sustainability Institute (SSI) explained:

> It is easy to concentrate on the short-term issues when developing scientific software. Deadlines for publications, collaboration with others and the demands of a daily routine all conspire to prevent proper planning. A software management plan can help to formalise a set of structures and goals that ensure research software is accessible and reusable in the short, medium and long term. It also helps researchers to consider whether third-party software to be used within a research project will be available, and supported, for the lifetime of the project. They can also give funders confidence that software they have funded survives beyond the funding period, that there is something to show for their investment (Jackson, 2015).

As Jackson noted, researchers are often under pressure and may not consider some of these issues until it is too late. Librarians, as part of their overall data management planning and curation support during the conceptualisation and creation stages, might help researchers to consider whether such a plan is required for their research and help them to complete the plan if necessary. In 2014, the SSI collaborated with the Digital Curation Centre (DCC) to develop a software management plan service (SSI, 2015) based on DMPonline.

Appraise and select stage

A key stage in the curation lifecycle is selecting what data must be kept beyond the life of the project. As a minimum, most funders expect sufficient data to be retained to enable validation of published research findings. However, data that cannot be reproduced or has potential longer term value may also be considered for longer term retention. Librarians could play a significant role in assisting researchers to assess what data must be retained to comply with funders' and publishers' expectations on reproducibility. Relevant CRediT taxonomy elements for the appraise and select stage of the curation lifecycle is shown in Table 3.

Recent reports have highlighted the worrying fact that a large proportion of published research findings are not reproducible. While reproducibility itself is not always a guarantee of integrity, an open and transparent approach is optimal as explained in a recent Nature Editorial:

> Reproducibility, rigour, transparency and independent verification are cornerstones of the scientific method. Of course, just because a result is reproducible does not make it right, and just because it is not reproducible does not make it wrong. A transparent and rigorous approach, however, will almost always shine a light on issues of reproducibility. This light ensures that science moves forward, through independent verifications as well as the course corrections that come from refutations and the objective examination of the resulting data (Nature, 2014).

However, as the size and the complexity of the data being generated and analysed in research activity increases, it is clear that replication will not always be feasible. In such cases, the early accessibility to the underlying data and tools used to produce published findings becomes even more crucial.

> In the era of Big Data and expensive science, it isn't always possible to replicate an experiment. However, it is possible to post the data and the computer software used to analyse it online, so that others can verify the results (Organizing Committee of the Future of the Statistical Sciences Workshop, 2014).

Early openness amongst peers can help to identify any problems or mistakes in the data and ensure that these are corrected. However, it can often be difficult for researchers to know how and when they should share their data. Indeed, many researchers are often confused by mandates relating to data sharing and feel unsure about what data can be shared and with whom it could be shared. The Expert Advisory Group on Data Access (EAGDA) found that "many researchers are unclear as to which research datasets should be made accessible to secondary researchers. This judgement should take into account the size, complexity and generalizable utility of the data generated" (EAGDA, 2015).

It is important to remember that funding bodies are not advocating that any researchers share data that will breach data protection legislation or infringe upon IPR associated with commercial research activity or emerging patents. Indeed, funding bodies expect researchers to employ exemptions to data sharing as appropriate. The Engineering and Physical Sciences Research Council expects that "researchers and research students have a general awareness of the regulatory environment and of the available exemptions which may be used, should the need arise, to justify the withholding of research data" (EPSRC Expectations, 2011). Librarians can help researchers to identify what data can be shared and also help identify the best mechanisms to share their data at various points over the research lifecycle – from the use of safe havens for access to sensitive data during the active phase of research through to deposit in subject specific or institutional data repositories for wider visibility beyond the life of the project.

Ultimately, data sharing at appropriate points in the research lifecycle enables peer review and helps improve the overall integrity of the published findings. Research integrity is important for researchers' career development and reputa-

Table 3: CRediT taxonomy elements for the appraise and select stage (Allen et al, 2014)

#4 validation	Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs
#7 resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools

Table 4: CRediT taxonomy elements for the creation stage (Allen et al, 2014)

#9 writing	Original draft: Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
#10 writing	Review and editing: Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre- or post-publication stages

tions, but equally vital for research organisations' overall credibility. In addition, research integrity helps to build confidence and trust in researchers and the findings which will be used to inform policy making. Finally, research integrity is crucial for ensuring that ongoing public investment is made available to support future research activity (Science Europe, 2015). However, not all research will be captured in a digital format. Many researchers work in a hybrid environment where analogue materials such as tissue samples or core samples form an integral part of the analysis. In some cases, funders may require that mechanisms for enabling access to non-digital materials are provided. As Markin (2015) pointed out "some divisions of the NSF require providing access only to digital data, while the ocean-sciences division of the NSF requires the sharing of sediment, core, and dredge samples as well". It is crucial that researchers consider how links to such analogue materials will be handled with regards to research integrity, reproducibility and validation of published findings. Librarians will play an important role in advising researchers on how to store, cite, and link the wide range of research materials collected and produced during their research activity.

Access, use and reuse

During the access, use and reuse stage of the curation lifecycle, researchers will be working to publish and share their research findings. This will involve understanding funding bodies' and publishers' requirements relating to timing and allowable embargo periods, as well as more practical aspects such as providing identifiers and statements indicating how underlying data might be accessed. Table 4 shows the relevant CRediT taxonomy elements for the creation stage.

Librarians have been instrumental in supporting open access with regards to research publications. Their experience with helping researchers to navigate through the complex maze of institutional polices, funding body requirements, and publishers' expectations will be an asset in assisting researchers to consider publishing data alongside their publications. As noted earlier, librarians may play a key role in supporting researchers to identify what data must – or should – be published and/or retained to enable validation of published findings or for broader reuse value. Data that is selected for publication and/or retention will require the application of a persistent identifier using services such as DataCite. The application of identifiers can offer benefits. On its website, DataCite states "Citable data become legitimate contributions to scholarly communication, paving the way for new metrics and publication models that recognise and reward data sharing" (DataCite, 2015).

Facilitating access to research outputs is an increasingly integral aspect of funded research activity. However, the effort required by librarians to enable effective access to research publications and related outputs is often underestimated by senior management. In addition to advising on stakeholder requirements, selecting an appropriate route for open access publication (i.e., green or gold), advising on the timing of deposit, for example, in relation to HEFCE's requirements for REF 2020 librarians will also need to deal with finding funds for Article Processing Charges (APCs), and ensuring that there are statements in published papers that specify how to access underlying data, and relevant software and code (RCUK Policy on Open Access, 2013). Additional effort will also be required to help researchers promote their work, to monitor usage, and to gather metrics. The Expert Advisory Group on Data Access (EAGDA) report stated:

> As data access activities are subsumed into general staff time and not specifically credited, it was felt by some interviewees that data access is supported primarily through general goodwill of staff, which is not a sustainable basis for maintaining access (EAGDA, 2015).

Librarians' contribution to the publication process is crucial in a number of ways and cannot be sustained without adequate resource allocation and recognition of this vital support role.

Conclusion

It is clear that the professions of both researchers and librarians are evolving as the research landscape changes. As demonstrated through the mapping to the CRediT taxonomy, there is no shortage in the number and range of opportunities available to librarians to become active players in supporting RDM and curation as part of the open science movement. Librarians have consistently demonstrated flexibility and willingness to embrace change and this attitude in no small part has been a crucial factor in the early successes we've seen in the UK.

We must ensure that emerging LIS professionals have the capacity to contribute effectively to this changing landscape and that there are sufficient opportunities for ongoing professional development that are endorsed by professional bodies and employers. A number of new educational programmes are being developed, e.g., the Library and Information Science Centre's MPhil in Digital Curation (University of Cape Town, 2015). However, we must bear in mind that librarians already have very heavy workloads and a wide range of work related responsibilities beyond open science. Indeed, librarians are, in many cases, active researchers in their own right who secure grant funding and publish. The full range of professional activities must be taken into account when developing new career specifications for librarians and – equally important – must also be recognised by employers.

Researchers and research support staff will need to acquire new skills to work effectively as the research landscape changes. To foster a real change in practice though, and to see these communities of practice embrace open science, we must ensure that the acquisition and application of these new skills are recognised and rewarded. The CRediT taxonomy could offer an effective means of describing the broad range and level of contributions from both researchers and support staff in a more unified way and help to inform future changes in research assessment exercises and related reward structures.

References

Allen L, Scott J, Brand A, Hlava M, Altman M. 2014. Publishing: Credit where credit is due. Nature 508. Available: http://www.nature.com/news/publishing-credit-where-credit-is-due-1.15033 (05 August 2015).

Association of Research Managers and Administrators (ARMA). 2011. Professional Development Framework. Available: https://www.arma.ac.uk/professional-development/PDF (05 August 2015).

Australian National Data Service (ANDS). Research Data Australia. Available: https://researchdata.ands.org.au/ (04 August 2015).

Cancer Research UK, Economic and Social Research Council, Medical Research Council, Wellcome Trust.2015. Funders' Response to the Expert Advisory Group on Data Access (EAGDA) Report on Data Access. Available: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@msh_peda/documents/web_document/wtp059353.pdf (05 August 2015).

Chartered Institute for Library & Information Professionals (CILIP). 2013. Professional knowledge and skills base (PKSB). Available: http://www.cilip.org.uk/cilip/jobs-and-careers/professional-knowledge-and-skills-base/what-professional-knowledge-and-skills (05 August 2015).

Collaboration to Clarify the Costs of Curation (4C). 2013. Available: http://4cproject.eu/ (05August 2015).

Collaboration to Clarify the Costs of Curation (4C). 2015. Curation Costs Exchange (CCEx) Available: Available: http://www.curationexchange.org/ (05August 2015).

CRediT. 2015. Available: http://credit.casrai.org/. (03 August 2015).

DataCite.2015. Available: https://www.datacite.org/ (05 August 2015).

Davidson, J., Corrall, S., Coulbourne, G., and Rauber, A. 2012. Education alignment.In: Aligning national approaches to digital preservation, N. McGovern and K. Skinner, K. (Eds.). Atlanta, GA: Educopia Institute Publications. 269-308. Available: http://educopia.org/publications/anadp. (05 August 2015).

Davidson. J and Jung, C. 2015. Proposed work plan for Research Data Alliance (RDA) Education and Training Interest Group. [Electronic mailing list, 9 February]. Available: https://rd-alliance.org/group/education-and-training-handling-research-data/post/proposed-work-data-science-curricula.html. (05 August 2015).

Digital Curation Centre (DCC). 2009. Data Asset Framework. Available: http://www.dcc.ac.uk/resources/tools/data-asset-framework.(05 August 2015).

Digital Curation Centre (DCC). 2010. Data Management Skills Support Initiative (DaMSSI). Available: http://www.dcc.ac.uk/training/damssi (05 August 2015).

Digital Curation Centre (DCC). 2011. Data Management Skills Support Initiative - Assessment, Benchmarking and Classification (DaMSSI-ABC) Available: http://www.dcc.ac.uk/training/damssi-abc (05 August 2015).

Digital Curation Centre (DCC). 2014. Five steps to decide what data to keep. Available: http://www.dcc.ac.uk/resources/how-guides/five-steps-decide-what-data-keep (05 August 2015).

Digital Curation Centre (DCC). 2015. DCC Curation Lifecycle Model. Available: http://www.dcc.ac.uk/resources/curation-lifecycle-model (05August 2015).

Digital Curation Centre (DCC). 2015. Available: http://www.dcc.ac.uk (05 August 2015).

Digital Curation Centre (DCC). 2015. DMPonline tool. Available: https://dmponline.dcc.ac.uk/projects (05 August 2015).

Economic and Social Research Council (ESRC). 2015. Data management planning for ESRC researchers. Available: http://ukdataservice.ac.uk/manage-data/plan/dmp-esrc (05 August 2015).

EDINA, University of Edinburgh. 2010. MANTRA Research Data Management Training. Available: http://datalib.edina.ac.uk/mantra/ (05 August 2015).

Engineering and Physical Sciences Research Council (EPSRC). 2011. EPSRC policy framework on research data. Available: https://www.epsrc.ac.uk/about/standards/researchdata/ (05 August 2015).

Engineering and Physical Sciences Research Council (EPSRC). 2011. Expectations; Availa-

ble: https://www.epsrc.ac.uk/about/standards/researchdata/expectations/ (05 August 2015).

Expert Advisory Group on Data Access (EAGDA). 2015. Report on governance of data access and annexes. Available: http://www.wellcome.ac.uk/stellent/groups/corporatesite/@msh_peda/documents/web_document/wtp059343.pdf (05 August 2015).

Higgins, S. 2008. The DCC curation lifecycle model. IJDC, 3 (1):134-140. doi:10.2218/ijdc.v3i1.48

Higher Education Funding Council for England (HEFCE). 2015. Policy for open access in the post-2014 Research Excellence Framework. Available: http://www.hefce.ac.uk/pubs/year/2014/201407/ (05 August 2015).

Higher Education Funding Council for England (HEFCE), the Scottish Funding Council (SFC), the Higher Education Funding Council for Wales (HEFCW) and the Department for Employment and Learning, Northern Ireland (DEL). 2008. Research Assessment Exercise (RAE). Available: http://www.rae.ac.uk/ . (05 August 2015).

Higher Education Funding Council for England (HEFCE), the Scottish Funding Council (SFC), the Higher Education Funding Council for Wales (HEFCW) and the Department for Employment and Learning, Northern Ireland (DEL). 2014. Research Excellence Framework (REF): Assessment framework and guidance on submissions. 2011. Available: http://www.ref.ac.uk/media/ref/content/pub/assessmentframeworkandguidanceonsubmissions/GOS%20including%20addendum.pdf (05 August 2015).

Hubbard, B. 2015. Getting the rights right - or, When policies collide!, UKSG Webinar, 19 May 2015. Available: http://www.uksg.org/webinars/oapolicies. (05 August 2015).

Jackson, M. 2015. Software management plan service prototype live. [blog, March 21 2015]. Available: http://www.dcc.ac.uk/blog/software-management-plan-service-prototype-live (05 August 2015).

Jisc. Research Data Discovery Service. Available: http://www.jisc.ac.uk/rd/projects/uk-research-data-discovery (05 August 2015).

Markin, K. 2015. Where Should You Keep Your Data? The Chronicle for Higher Education. June 23. Available: http://chronicle.com/article/Where-Should-You-Keep-Your/231065 (05 August 2015).

National Oceanic and Atmospheric Administration (NOAA). National Centers for Environmental Information (NCEI). Available: http://www.ncdc.noaa.gov/data-access (05 August 2015).

Nature. 2014. Journals unite for reproducibility. Editorial in Nature Special Issue: Challenges in irreproducible research. Available: http://www.nature.com/nature/focus/reproducibility/index.html. (04 August 2015).

Organizing Committee of the Future of the Statistical Sciences Workshop. Statistics and Science: A Report of the London Workshop on the Future of the Statistical Sciences. Available: http://www.socialscienceespace.com/2014/07/statistics-crisis-of-reproducibility/ (03 August 2015).

Project CRediT. 2015. Available: http://credit.casrai.org/. (03 August 2015).

Research Councils UK (RCUK). 2011. Common Principles on Research Data Policy. Available: http://www.rcuk.ac.uk/research/datapolicy/ (04 August 2015).

Research Councils UK (RCUK) Policy on Open Access. 2013. Available: http://www.rcuk.ac.uk/research/openaccess/policy/ [accessed August 4, 2015] (04 August 2015).

Ryan. B. 2013. RCUK responses to questions raised at the DCC/RDMF special event on funding for research data management held at Aston University on 25th April 2013. Available: http://blogs.rcuk.ac.uk/files/2013/07/RCUK-Responses-to-DCC-RDMF-Funder-Questions-.pdf (04 August 2015).

Science Europe. Working Group on Research Integrity. 2015. Seven reasons to care about in-

tegrity in research. Available: http://www.scienceeurope.org/uploads/PublicDocumentsAndSpeeches/WGs_docs/20150617_Seven%20Reasons_web2_Final.pdf (04 August 2015).

Software Sustainability Institute (SSI). 2015. Software Management Service Prototype. Available: https://ssi-dev.epcc.ed.ac.uk/ (05 August 2015).

UK Data Archive. 2013. Data Management Costing Tool and checklist. Available: http://www.data-archive.ac.uk/create-manage/planning-for-sharing/costing. (03 August 2015).

University of Cape Town. Library and Information Studies Centre (LISC), 2015. MPhil with a specialisation in Digital Curation. Available: http://www.lisc.uct.ac.za/digitalcuration_mphil (August 5 2015).

Vitae. 2012. Information literacy lens on the Vitae Researcher Development Framework using the SCONUL Seven Pillars of Information Literacy. Available: https://www.vitae.ac.uk/vitae-publications/rdf-related/information-literacy-lens-on-the-vitae-researcher-development-framework-rdf-apr-2012.pdf. (04 August 2015).