

10th Italian Research Conference on Digital Libraries, IRCDL 2014

Research Data Management in the curriculum: an interdisciplinary approach

Anna Maria Tammaro^{a*}, Vittore Casarosa^b

^aUniversity of Parma, Viale Usberti 181a, 43125 Parma

^bISTI-CNR, Via Moruzzi 1, 56124 Pisa, Italy

Abstract

Research Data Management is broadly understood as collecting, analyzing, publishing, reanalyzing, critiquing, and reusing data. The increase of digital content in the broad areas of Institutional and domain specific Repositories, Libraries, Archives and Museums and the increased interest in the sharing and preservation of "research data" have triggered the emergence of new roles such as Data Curator. The paper refers about the on-going investigation of current data curator education and training programs with regard to the role of information professionals and/or data scientists in the research lifecycle. The investigation has been based on a series of workshops and events discussing the concerns of researchers and teachers about digital library and digital curation. A first list of competencies and skills at technical and operational level that professionals should have, has been evidenced. The theoretical framework and structure of educational programmes should have sufficient flexibility to accommodate the needs of various groups of specialists.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Peer-review under responsibility of the Scientific Committee of IRCDL 2014

Keywords: Research Data Management; Digital Library Education; Data Curator;

1. Introduction

Research Data Management is the umbrella term of the movement to make scientific research data and dissemination material accessible to all levels of an inquiring society, amateur or professional. It encompasses practices such as open publishing of research, campaigning for open access, encouraging scientists to practice open

* Corresponding author. Tel.: +390521905703.

E-mail address: annamaria.tammaro@unipr.it

notebook science, and generally making it easier to publish and communicate scientific knowledge. Given the scope of activities required, Research Data Management may be performed by tenure-track scholars, alternative academics, librarians, software developers, students both undergraduate and graduate, as well as interested members of the public. Research Data Management and in particular data curation address the challenge of maintaining digital information that is produced in the course of research in a manner that preserves its meaning and usefulness as a potential input for further research.

The increase of digital content in the broad areas of Institutional and domain specific Repositories, Libraries, Archives and Museums and the increased interest in the sharing and preservation of research data have triggered the emergence of new competencies and skills such as the role of Digital Curator. Researchers need better services that provide more sophisticated access controls, collaboration and can support the volume of data generated. This can have important implications for the design of curation systems and curation education programs. The main questions of interest are:

- More and more the professions of Librarian, Archivist and Museum curator seem to converge into the single role of an "Information professional". Is there a common core of knowledge/skills that could be taught? What is it? Should there be a Data Librarian, or a Data Archivist, or a Data Museum curator?^{1,2} Or is this a new role to be invented from scratch?
- In addition to the traditional Librarian, Archivist and Museum curator curricula, what are the specific knowledge/skills that need to be taught in each discipline? Or should the responsibility of curating research data be given to the "data producers", i.e. the researchers themselves?
- Is the role/skills of a "Data Curator/Custodian" in the research/scientific domains the same as the role/skills of a Librarian/Archivist/Museum curator in the Cultural Heritage domain?

With the increase of digital content to be made available by Libraries, Archives and Museums, the interest of the authors for an interdisciplinary approach to the emerging role of "information professionals" began in 2005, which already appears as a long time ago.

1.1. Methodology

This short paper reports about an on-going investigation of current education and training programs in research data curation, with regard to possible competencies gap, and the role of information professionals and data scientists in the research data management lifecycle. The investigation is based on the ongoing interdisciplinary discussion held throughout a series of workshops and events discussing the concerns about digital library and digital curation of researchers and teachers. As a team with different backgrounds (Humanities and Computer Science, Education and Research), the search for a common understanding of the educational needs of the information professionals has been brought forward through a long series of workshops, seminars and panels^{5, 6, 7, 10, 17, 19}. As already mentioned, the first event was a workshop held in 2005 at the University of Parma, up to the last two workshops held in 2013, in connection with the iConference 2013 and the TPDL 2013 Conference. A detailed description of the different events and an analysis of their results can be found in^{16, 18, 20}.

1.2. Literature review

For the purpose of this paper, Research Data Management is best described as life cycle data management; it encompasses a spectrum of activities ranging from research data management planning at the project inception stage; through collection of data as part of the research process; through the identification, processing, and accession of data sets; and, finally, to the archival preservation and sharing of data in an appropriate repository. Those activities add value and knowledge to the collections, and the added value is usually given by the curator or manager of the institution. In this way Research Data Management is distinct from digital curation or digital preservation (though preservation represents an important part of curation) as it encourages stronger connections between library, technology and research scientists. In the literature, very often research data management, data curation, digital library, digital preservation are used interchangeably.

As it already happened for Digital Libraries, the United States are at the forefront for what concerns education in data curation and Research Data Management. For example, the Institute of Museum and Library Services', through

the Laura Bush 21st Century Librarian Program¹², has been supporting a significant range of projects in “education for the next generation of librarians, faculty, and library leaders”. At the University of North Carolina at Chapel Hill, a matrix of skills and functions has been developed¹³, describing 24 functional areas and 4 meta-level functions. The University of Illinois at Champaign-Urbana, the University of Tennessee at Knoxville and the National Center for Atmospheric Research have promoted an alliance called “Data Curation Education in Research Centers”⁸, aimed at developing a sustainable and transferable model for educating LIS masters and doctoral students in data curation, through field experiences in research and data centers.

In Italy, the discussion about Research Data Management has been limited, and has been focussing mostly on the preservation aspect of the data life-cycle. Several research organizations have been involved in activities and in European Projects related to digital preservation. The *Fondazione Rinascimento Digitale*, in 2012 organized an International Conference on “Trusted digital repositories and trusted professionals”, focussing on the shortage of competencies and skills for digital preservation and curation. Some Italian initiatives were presented, such as the project on “Digital Stacks”, developed at the Central National Library of Florence for the legal deposit and preservation of Italian digital publications³, or the good practices for digital preservation in Italian public administrations¹¹ or one of the first attempts to preserve research data produced in Italian universities⁴.

In 2013 DigCurV⁹ (a European project in which *Fondazione Rinascimento Digitale* was a partner) organized an international conference on “Framing the digital curation curriculum”, where the main objective was “to promote discussion and consensus building amongst stakeholders about the main criteria and requirements necessary to develop training courses for professionals in digital curation in the cultural heritage sector”. There also some interesting Italian activities were presented. The University of Parma²⁰ is an active partner in DILL (Digital Libraries Learning, an international master until recently funded by the EU Erasmus Mundus program), where the aim is to provide both a theoretical and conceptual foundation on Digital Libraries and some more practical experience (through courses and internships) on the technologies needed for the management of digital data. The University of Torino²¹ is launching a “training course” (rather than a Master) whose main aim is to provide “a broad perspective of the role of the digital curator”.

2. First findings

A first list of required competences and skills at technical and operational level that professionals have been evidenced appears in¹⁷, based on a Delphi study developed in the frame of a DILL Master thesis¹⁴. A literature review has led to the development of a conceptual model as depicted in Figure 1.

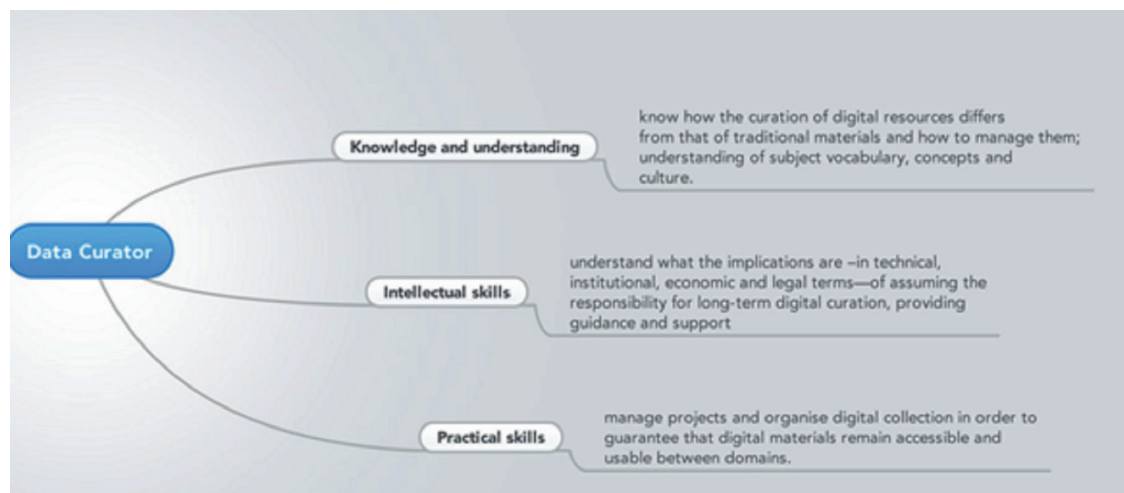


Figure 1 A conceptual view of data curator competences

The competences of the data curator are to be described as: knowledge and understanding, intellectual skills (ability to apply knowledge in different contexts) and practical skills (ability to manage projects and organizations).

- *Knowledge and understanding*: know how the curation of digital resources differs from that of traditional materials and how to manage them; understanding of subject vocabulary, concepts and culture.
- *Intellectual skills*: understand what the implications are—in technical, institutional, economic and legal terms—of assuming the responsibility for long-term digital curation, providing guidance and support
- *Practical skills*: manage projects and organise digital collection in order to guarantee that digital materials remain accessible and usable between domains.

The on-going investigation has evidenced competency gaps in all the three areas, related to organizational and policy aspects, and also to issues related to the data itself.

3. Conclusions

From all these activities, it appears clearly that Research Data Management and curation is not merely a technical/engineering issue, but it involves many different stakeholders and data curators, who need to be aware of the organisational context, the disciplinary methodology and the economic, social and political issues; the digital curator should be involved in the entire life-cycle of a digital resource, from its creation to its preservation for “future generations”. It also appears that regardless of the origin and the intended fruition of the digital resource, large segments of its life cycle are more or less the same in each of the three traditional disciplines of libraries, archives and museums. Of course, given the different focus and mission of the three disciplines, the value adding and the access portions of the life cycle will remain different, but we believe that further collaboration for the development of a common curriculum in digital curation can be built upon the many similarities over the entire life cycle.

At the same time, as it could be expected, the three research questions mentioned at the beginning remain largely unanswered. In this type of methodology (workshops and panels) firm conclusions are never reached; it is rather a continuous (slow) process in which at each step some more insight into the issues is acquired, but at the same time new questions are discovered. In general terms it can be said that over time, for what concern educational needs, there has been a shift from technology knowledge (which today is often considered an acquired pre-requisite) to knowledge and understanding of the workflow and the entire life cycle of the (digital) data.

The convergence of education of information professional in archives, museums, and libraries remains one of the main topics to be addressed in education. While evidence of such convergence has been found in North America and in some other places internationally, clearly there is considerable additional research and discussion that is needed to determine the most effective way to proceed in order to realize such convergence, especially in Italy. The discussion has to be continued, as the focus on competencies needed by information professionals in these three fields, as they relate to the curation and management of research data collections is clearly pushing education towards a multidisciplinary background of the professional.

From the examination of existing education and training programs and the results of the workshops, there is a clear need for data curation to move beyond passively providing technology for storage and preservation, to embrace the changes in scholarly production that emerging technologies have brought, to make sense of the digital curator role as a whole. The theoretical framework and structure of educational programs should have sufficient flexibility to accommodate the needs of the various groups of specialists.

The increased interest for “open research” and the availability of research data is changing the traditional scholarly research workflow, which was basically inspiration, exploration, discovery, followed by formulation, research design, data collection, followed by analysis and interpretation and finally concluding with documentation, publication and dissemination. The research workflow is now evolving towards a global cyber-infrastructure providing preservation and access to research data and research results, and providing discipline-specific tools. In this model the role of the Memory Institutions should evolve from one of holders and providers of knowledge resources to one of being an active partner in the research process, where the information professionals provide tools and expertise to the “research data producers”, in order to expedite and improve research and scholarship.

References

1. Borgman C. (2010) "Why data matters to librarians – and how to educate the next generation. The Changing Role of Libraries in Support of Research Data Activities: A Public Symposium", 2010, http://sites.nationalacademies.org/PGA/brdi/PGA_056901
2. Borgman C. (2012) IS289 "Data, Data Practices, and Data Curation". Los Angeles: Department of Information Studies, University of California, 2012, http://polaris.gseis.ucla.edu/cborgman/Chriss_Site/Courses.html
3. Bergamin G. (2012), "Magazzini digitali = The Digital Stacks", Int. Conference Trusted Digital Repositories & Trusted Professionals, 11-12 December 2012, Firenze, <http://93.63.166.138:8080/dspace/handle/2012/119> (last visited Dec. 29th, 2013)
4. Boschini M, Cortese C. (2012), "The road (and the roadmap) for building Trusted Digital Repositories within an Interuniversity Consortium", Int. Conference Trusted Digital Repositories & Trusted Professionals, 11-12 December 2012, Firenze, <http://93.63.166.138:8080/dspace/handle/2012/108> (last visited Dec. 29th, 2013)
5. Casarosa V, Cousins J, Tammamro AM, Ioannidis I. (2008), "The Web Versus Digital Libraries: Time to Revisit This Once Hot Topic", in Birte Christensen-Dalsgaard, Donatella Castelli, Bolette Ammitzbøll Jurik, Joan Lippincott, editors, Research and Advanced Technology for Digital Libraries, 12th European Conference, ECDL 2008, Aarhus, Denmark, September 14-19, 2008. Proceedings. Volume 5173 of Lecture Notes in Computer Science, pages 383-384, Springer, 2008, http://dx.doi.org/10.1007/978-3-540-87599-4_39, <http://www.ecdl2008.org/panels/>
6. Casarosa V, Castelli D, Tammamro AM. (2011), Report on the Workshop "Linking Research and Education in Digital Libraries, D-Lib Magazine, Vol. 17, No. 11/12, November/December 2011, doi:10.1045/november2011-casarosa, <http://www.dl.org.eu/index.php/dl-org-events/tpdl-workshop-linking-research-educationin-dlsn-in-dls>
7. DataCur2013 (2013), Workshop "Moving beyond technology: iSchools and education in data curation. Is Data Curator a new role?" (in connection with TPDL 2013, 23-26 September, La Valletta, Malta) <http://datacur.isti.cnr.it/index.html>, (last visited Dec. 16th, 2013)
8. DCERC (2013), Data Curation Education in Research Centers, <http://www.sis.utk.edu/dcerc>, (last visited Dec. 29th, 2013); <http://www.lis.illinois.edu/research/projects/dcerc>, (last visited Dec. 29th, 2013)
9. DigiCurV (2010), "Digital Curator Vocational Education Europe", <http://www.digcurededucation.org/index.php/eng> (last visited Dec. 29th, 2013)
10. DL.org and DILL (2010), Workshop Research and Education for Digital Libraries, Parma, November 2010, <http://www.dl.org.eu/index.php/dl-org-events/research-education-indigital-libraries>
11. Guercio, M. (2012), "The Italian case: legal framework and good practices for digital preservation", Int. Conference Trusted Digital Repositories & Trusted Professionals, 11-12 December 2012, Firenze, <http://93.63.166.138:8080/dspace/handle/2012/115> (last visited Dec. 29th, 2013)
12. IMLS (2013), "Institute of Museum and Library Services", Laura Bush 21st Century Librarian Program, <http://www.imls.gov/applicants/detail.aspx?GrantId=9> (last visited Dec. 29th, 2013)
13. Lee CA, Tibbo H. (2011). Where's the archivist in digital curation? Exploring the possibilities through a matrix of knowledge and skills. *Archivaria*, 72, 123-168.
14. Madrid M. (2011), "A Study of Digital Curator Competences: A survey of experts", Master Thesis, DILL International Master Digital Library Learning, University of Parma
15. Tammamro AM. (2006), "IT profiles and curricula for digital libraries in Europe", Int. Conference Libraries in the Digital Age (LIDA 2006), 29th May-3rd June 2006, Dubrovnik, https://dspace.unipr.cilea.it/bitstream/1889/1185/1/Tammamro_LIDA_2006.pdf (last visited Dec. 29th, 2013)
16. Tammamro AM, Madrid M, Casarosa V. (2012), Digital curators' education: professional identity vs. convergence of LAM (Libraries, Archives, Museums), IRCDL 2012, 9-10 February, Bari, Italy, in Communications in Computer and Information Science Volume 354, 2013, pp 184-194, http://link.springer.com/chapter/10.1007%2F978-3-642-35834-0_19, (last visited Dec. 29th, 2013)
17. Tammamro AM, Casarosa V, Borgman C, Silipigni Connaway L, Castelli D, Radford M. (panelists) (2012), "Can Research help Education in Digital Libraries?", Int. Conference Libraries in the Digital Age (LIDA 2012), 18-22 June 2012, Zadar, <http://ozk.unizd.hr/lida/2012/program/> (last visited Dec. 29th, 2013)
18. Tammamro AM, Casarosa V, Castelli D. (2013), "Closing the gap: interdisciplinary perspectives on research and education for digital libraries", IRCDL 2013, Rome, Italy, to appear in Communications in Computer and Information Science, Springer, 2013.
19. Tammamro AM, Casarosa V, Ross S, Moulaison H, Weech T, Lugya F. (2013). iSchools building on the strengths found in the convergence of librarianship, archival, and museum studies to improve the education of managing digital collections. iConference 2013 Proceedings (pp. 1024-1025), <http://hdl.handle.net/2142/42495>
20. Tammamro AM, "Integrating Digital Curation in a Digital Library curriculum: the International Master DILL case study", Final DigCurV conference Framing the Digital Curation Curriculum, 6-7 May 2013, Firenze, <http://www.digcurededucation.org/eng/Resources/DigCurV-2013-proceedings/Tammamro-paper20> (last visited Dec. 29th, 2013)
21. Vivarelli M, Cassella M, Valacchi F, "The digital curator between continuity and change: developing a training course at the University of Turin", Final DigCurV conference Framing the Digital Curation Curriculum, 6-7 May 2013, Firenze, <http://www.digcurededucation.org/eng/Resources/DigCurV-2013-proceedings/Vivarelli-paper22> (last visited Dec. 29th, 2013)