



## Journal of Documentation

The conceptual landscape of digital curation

Alex H. Poole

### Article information:

To cite this document:

Alex H. Poole, (2016), "The conceptual landscape of digital curation", Journal of Documentation, Vol. 72 Iss 5 pp. 961 - 986

Permanent link to this document:

<http://dx.doi.org/10.1108/JD-10-2015-0123>

Downloaded on: 10 November 2016, At: 20:24 (PT)

References: this document contains references to 201 other documents.

To copy this document: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)

The fulltext of this document has been downloaded 337 times since 2016\*

### Users who downloaded this article also downloaded:

(2016), "Information in the knowledge acquisition process", Journal of Documentation, Vol. 72 Iss 5 pp. 930-960 <http://dx.doi.org/10.1108/JD-10-2015-0122>

(2016), "Pictorial metaphors for information", Journal of Documentation, Vol. 72 Iss 5 pp. 794-812 <http://dx.doi.org/10.1108/JD-07-2015-0080>

Access to this document was granted through an Emerald subscription provided by emerald-srm:563821 []

### For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit [www.emeraldinsight.com/authors](http://www.emeraldinsight.com/authors) for more information.

### About Emerald [www.emeraldinsight.com](http://www.emeraldinsight.com)

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

\*Related content and download information correct at time of download.

# The conceptual landscape of digital curation

Landscape of  
digital  
curation

Alex H. Poole

*Drexel University, Philadelphia, Pennsylvania, USA*

961

Received 3 October 2015  
Revised 6 April 2016  
Accepted 10 April 2016

## Abstract

**Purpose** – The purpose of this paper is to define and describe digital curation, an emerging field of theory and practice in the information professions that embraces digital preservation, data curation, and management of information assets over their lifecycle. It dissects key issues and debates in the area while arguing that digital curation is a vital strategy for dealing with the so-called data deluge.

**Design/methodology/approach** – This paper explores digital curation's potential to provide an improved return on investment in data work.

**Findings** – A vital counterweight to the problem of data loss, digital curation also adds value to trusted data assets for current and future use. This paper unpacks data, the research enterprise, the roles and responsibilities of digital curation professionals, the data lifecycle, metadata, sharing and reuse, scholarly communication (cyberscholarship, publication and citation, and rights), infrastructure (archives, centers, libraries, and institutional repositories), and overarching issues (standards, governance and policy, planning and data management plans, risk management, evaluation, and metrics, sustainability, and outreach).

**Originality/value** – A critical discussion that focusses on North America and the UK, this paper synthesizes previous findings and conclusions in the area of digital curation. It has value for digital curation professionals and researchers as well as students in library and information science who may deal with data in the future. This paper helps potential stakeholders understand the intellectual and practical framework and the importance of digital curation in adding value to scholarly (science, social science, and humanities) and other types of data. This paper suggests the need for further empirical research, not only in exploring the actual sharing and reuse practices of various sectors, disciplines, and domains, but also in considering the the data lifecycle, the potential role of archivists, funding and sustainability, outreach and awareness-raising, and metrics.

**Keywords** Collaboration, Information management, Data, Data handling, Interdisciplinarity, Data management, Digital curation, Data curation

**Paper type** Conceptual paper

We swim in a sea of data [...] and the sea level is rising rapidly (Anderson and Rainie, 2012).

Virtually no one in academia perceives that they have a professional responsibility or mandate for research data management functions (Council on Library and Information Resources, 2013, p. 6).

## Introduction

“Contemplating the digital universe is a little like contemplating Avogadro’s number,” claim Gantz *et al.* (2008). “It’s big. Bigger than anything we can touch, feel, or see, and thus impossible to understand in context” (p. 3). IBM concludes, “Every day, we create 2.5 quintillion bytes of data—so much that 90% of the data in the world today has been created in the last two years alone[1].” Researchers grapple with “a tsunami of information that paradoxically feeds the growing scientific output while simultaneously crushing researchers with its weight” (Haendel *et al.*, 2012, p. 1). The data deluge is truly upon us.

That data deluge presents unprecedented challenges in preserving digital assets across all sectors of society, whether organizational, technological, legal, cultural, or business. Rothenberg (1995) opts for an apothegm: “Digital objects last forever – or



five years, whichever comes first” (p. 42). Other scholars worry about the potential loss of important data, identifying “A Public Trust at Risk” (Heritage Preservation, 2005), data’s “shameful neglect” (*Nature*, 2009), or the specter of a “digital dark age” (Bollacker, 2010). Technical obsolescence or fragility, lack of resources, ignorance of good practices, uncertainty over appropriate infrastructure – all constitute serious risks to data (Harvey, 2010). Digital curation tackles these risks.

This paper concentrates on foundational Anglo-American digital curation research and practice. It first defines digital curation, discusses its activities, and sets forth the roles and responsibilities of digital curators. Second, it unpacks data and their role in the research enterprise. Along these lines, it argues for the vital importance of the data lifecycle and of metadata to digital curation. Third, sharing and reuse of data are discussed; digital curation facilitates both practices. Fourth, this paper addresses researcher behaviors and the ways in which disciplinarity, communities of practice, and collaboration shape the sharing and reuse of data. Fifth, digital curation’s impact upon scholarly communication, namely, on cyberscholarship, publication and citation, and rights, is diagramed. Sixth, digital curation needs to be embedded in institutional and scholarly infrastructure: archives, centers, libraries, and institutional repositories (IRs) are key components of that infrastructure. Seventh, standards, governance and policy, planning and data management plans (DMPs), risk management, evaluation, and metrics, sustainability, and outreach represent overarching concerns for digital curation stakeholders[2]. Finally, directions for future research are suggested.

### Digital curation

Digital curation centers on “maintaining and adding value to a trusted body of digital information for current and future use[3].” First used in 2001, the term embraces digital preservation, data curation, and the management of assets over their lifecycle (Lee and Tibbo, 2007; Yakel, 2007).

Digital curation bridges research, practice, and training across nations, disciplines, institutions, repositories, and data formats (Gold, 2010; Ray, 2009). Given the diversity of its stakeholders and of the environments in which it is conducted, it potentially involves anyone who interacts with digital information during its lifecycle (Dallas, 2007; Winget *et al.*, 2009).

The National Research Council of the National Academies (2015) concludes, “The field has grown from practices hardly recognized as curation per se [...] to international consortia engaged in defining shared norms and standards for digital curation” (p. 17). Researchers, moreover, increasingly recognize its salience. One recent study found that 90 percent of doctoral supervisors, doctoral holders, and students view digital curation as moderately or extremely important (Abbott, 2015). In spite of the long-term importance of digital curation, however, researchers often postpone it (Rusbridge, 2007). This is not necessarily surprising, as they face exigent challenges: a lack of standards, of a common vocabulary, of authority control(s), of appropriate hardware and software, and of storage space (Latham and Poe, 2012; Pryor, 2013).

### Digital curators

Digital curation implicates roles and responsibilities that meld library and information science (LIS) and non-LIS domain skills (Vivarelli *et al.*, 2013). Notwithstanding technical skills, digital curators require so-called soft skills, namely project management, negotiation, team-building, and collaborative problem solving (Harvey, 2010; Swan and Brown, 2008). Educational programs should therefore cultivate “professional allrounders like a Swiss army knife” (Osswald, 2013).

The role of digital curation professional is hardly discrete and defined; rather, diverse persons undertake variegated tasks in equally variegated settings (National Research Council of the National Academies, 2015).

Digital curation work may seem mundane or even invisible (Osswald, 2013). Indeed, neither the demand nor the impact of curation activities can always easily be measured, much less communicated (Abrams, 2014; National Research Council of the National Academies, 2015). But in advising creators as well as users, ensuring long-term access, facilitating discovery, retrieval, use, and reuse, promoting interoperability, and helping users maximize the usefulness of the curated content, digital curators perform essential tasks (Harvey, 2010).

The Education Advisory Board (2014) pinpoints a dearth of qualified professionals but an increased demand for such professionals in both the public and the private spheres. National employer demand for digital curation professionals increased by more than 50 percent between H1 of 2010 and H1 of 2013 and by 10 percent more between H2 of 2013 and H1 of 2014 (Education Advisory Board, 2014).

## Data

### *Defining data*

Unprecedentedly vast in scale and scope, data undergird the scientific record and enable the production of new knowledge (National Academy of Science, 2009; Pryor, 2012; Rusbridge, 2007). Despite this lofty station, the very notion of “data” remains vexatious, seemingly immune even to precise definition. For example, the National Science Board (2005) defines data as “any information that can be stored in digital form, including text, numbers, images, video or movies, audio, software, algorithms, equations, animations, models, simulations, etc.” (p. 9). Data can be framed in even more protean ways: they, “like beauty, exist in the eye of the beholder,” Borgman *et al.* (2012) insist (p. 517). Borgman (2015) similarly expounds, “Data can be many things to many people, all at the same time” (p. xvii).

Grappling with this malleability, stakeholders may define data through example, labeling facts, numbers, or symbols “data.” Similarly, they may resort to operational definitions such as those found in the open archival information system (OAIS) model and the data documentation initiative (DDI) (Borgman, 2015). The National Science Board (2005) adds some clarity by adumbrating a taxonomy based on three facets: the nature of the data, including their format (images, software, or models, for example) or their origins (observational, computational, experimental, or derivative); the data’s degree of processing; and their potential reproducibility. But overall, definitional quandaries persist. This inability to achieve consensus hamstrings efforts to optimize DMPs, open data policies, and curation overall (Borgman, 2015).

### *The data lifecycle and lifecycle models*

There is an unprecedented need not only to address the entire lifespan of digital content but to do so as early as possible in the research process (Ball, 2012; Tibbo, 2015). Digital content’s lifecycle embraces appraisal and ingest, classification, indexing, cataloging, and authority management, enhancement, presentation, publication, and dissemination, user experience and modeling, preservation, and repository management (Constantopoulos and Dallas, 2010). Data are generally most vulnerable at transition points in the information lifecycle (Corti *et al.*, 2014).

Lifecycle models describe the ways in which stakeholders preserve as well as add value to data (Pryor, 2012). Models allow stakeholders to map their work progress, to

discern vulnerabilities, to encourage documentation, to develop standards, and to identify tools and services; thus they help preserve data's authenticity, reliability, and usability (Harvey, 2010; Higgins, 2008).

Eight lifecycle models merit consideration: the DIGITAL CURATION CENTER CURATION LIFECYCLE MODEL (Higgins, 2008), the I2S2 Idealized Scientific Research Activity Lifecycle Model[4], the DDI Combined Life Cycle Model[5], ANDS Data Sharing Verbs[6], the DataONE Data Lifecycle[7], the Research360 Institutional Research Lifecycle[8], the Capability Maturity Model for Scientific Data Management [9], and the UK Data Archive Data Lifecycle[10]. Finally, the OAIS model broke new ground in the late 1990s and became ISO Standard 14721: 2003. It does not, however, constitute a full-fledged lifecycle model; it also neglects to specify guidelines for creating or (re)using data (Lee, 2005, 2009).

In developing or following an appropriate lifecycle model, designers should consider issues such as scope (what services will be offered and to what audiences), best practices or community standards or both, and the model's ability to represent work practices (Carlson, 2014). Lifecycle models are vital resources for digital curation work. So too are metadata.

### *Metadata*

Metadata deal with data attributes "that describe, provide context, indicate the quality, or document other object (or data) characteristics" (Greenberg, 2005, p. 20). Supporting nearly all of the steps in the digital curation lifecycle, they are tantamount in importance to data themselves (Levine, 2014; Riley, 2014). But one recent study underscores researchers' unfamiliarity with creating or documenting metadata; another discerns that researchers put nominal effort into metadata creation because they cannot predict the needs of future reusers (Akers and Doty, 2013; Wallis *et al.*, 2013). Edwards *et al.* (2011) lament researchers' recalcitrance in recording even basic metadata.

The metadata's benefits include not only a controlled vocabulary, but also information on related objects, on intellectual property, on user information, on versioning, on integrity checks, and on preservation (Harvey, 2010). Stakeholders consider abundant structured metadata a "holy grail" with respect to sharing and reuse (Edwards *et al.*, 2011, p. 672).

## **Sharing and reuse**

### *Sharing*

The sharing of data allows scholars to provide descriptive or historical context, to detect fraud or to address disputes, to reproduce or to verify research findings, to break new ground in research methodology, to make findings generated by publicly funded research available, to enable meta-analysis, to increase citation, to reduce loss, to enrich pedagogy, to generate new knowledge from familiar data and thus to leverage investments, to avoid the costs of re-collecting data, and to foster economic development (Borgman, 2012, 2015; Corti *et al.*, 2014; Faniel and Zimmerman, 2011; Heidorn, 2008; Lyon, 2009; McLure, *et al.*, 2014; Organisation for Economic Co-Operation and Development, 2007; Parsons and Duerr, 2005; Ray, 2014; Tenopir, *et al.*, 2011; Whitlock, 2011; Zimmerman, 2008).

Researchers who consider sharing must know which data can be shared and why, with whom, under what conditions, and to what effect: discipline, age, research as opposed to teaching focus, and geographical region comprise important

considerations (Borgman, 2012; Tenopir *et al.*, 2011). Researchers can share data through deposit in a data center, archive, bank, or institutional repository (IR), through submission to a journal as part of an article, through discrete publication, through mounting them online, and through peer exchange (Akers and Doty, 2013; Van den Eynden, *et al.*, 2010; Wallis, *et al.*, 2013).

Direct contact among researchers and other professionals remains the most important precondition for collaboration (Kroll and Forsman, 2010). Trust nurtured through personal relationships usually undergirds sharing (Akers and Doty, 2013; Akmon *et al.*, 2011; Cragin *et al.*, 2010; Duerr *et al.*, 2004; Faniel and Jacobsen, 2010; Kroll and Forsman, 2010; Pryor, 2009; Sayogo and Pardo, 2013; Tenopir *et al.*, 2015; Wallis *et al.*, 2013; Zimmerman, 2008). Private communication therefore plays an integral role: researchers not only can convey content, but also can provide documentation about the data's attributes and their applicability (Borgman, 2015). It also renders tacit knowledge more transferrable (Birnholtz and Bietz, 2003; Polanyi, 1966; Wallis *et al.*, 2013). On the other hand, private communication fails to enhance discoverability, usability, or longevity (Borgman, 2015).

Even if suitable data are available for sharing and reuse, disincentives persist. Edwards *et al.* (2011) insist, "Every moment of data across an interface comes at some cost in time, energy, and human attention" – i.e. "data friction" (p. 669). Six more granular concerns also apply. First, researchers may not believe their data can be useful to others (Steinhart, *et al.*, 2012). Prior to sharing, moreover, data must be gathered, formatted, structured, and documented in ways that accommodate particular domains and disciplines – a resource-intensive process (Akers and Doty, 2013; Corti *et al.*, 2014; Haendel, *et al.*, 2012; Ray, 2014; Strasser, 2015; Van den Eynden, *et al.*, 2010; Wallis *et al.*, 2013). Second, researchers fear they will not receive credit for sharing (Akers and Doty, 2013; Tenopir, *et al.*, 2011; Wallis, *et al.*, 2013). Third, researchers want an embargo period that allows them to analyze or otherwise use their data (Wallis *et al.*, 2013). For example, Cragin *et al.* (2010) found that 40 percent of respondents resisted sharing to an embargo period unless that sharing was circumscribed to collaborators or associates. In another study, although 95 percent of respondents expressed willingness to share their data, more than two-thirds (68 percent) wanted to wait at least six months after analysis to do so (Steinhart *et al.*, 2012). Fourth, researchers worry about misuse, primarily about misinterpretation, but also about intellectual property (Akers and Doty, 2013). One study determined that if their data were misused, researchers expressed less willingness to share data subsequently (Cragin *et al.*, 2010). Fifth, ethical concerns such as confidentiality or privacy can discourage sharing (Akers and Doty, 2013; Borgman, 2012; Corti *et al.*, 2015; Cragin *et al.*, 2010; Pryor, 2009; Steinhart *et al.*, 2012). Sixth, legal and rights issues can militate against sharing (Steinhart *et al.*, 2012). Not to be overlooked, finally, accessible data are not ipso facto usable data (Akmon, *et al.*, 2011; Tenopir, *et al.*, 2011; Wallis, *et al.*, 2013).

Discerning data's potential reuse value and the resources needed to ensure fit-for-purpose remains the litmus test for sharing (Palmer *et al.*, 2011). Most problematic, purported willingness to share may not in fact lead to data release, much less reuse (Borgman, 2015) (Scaramozzino *et al.*, 2012; Wallis *et al.*, 2013). Small wonder sharing seems a concept "almost byzantine" (Carlson *et al.*, 2011, p. 647) and is "maddeningly difficult" (Edwards *et al.*, 2011, p. 670).

### *Reuse*

Reuse may include migrating, enhancing, aggregating, or reanalyzing data (McLeod *et al.*, 2013). But reusing shared data erects further hurdles. First, the provenance of the

data set may be ambiguous (Zimmerman, 2008). Second, optimal contextualization and documentation relies upon understanding a discipline's, subdiscipline, or domain's history and its research community (Carlson and Anderson, 2007). Indeed, researchers often defer to disciplinary norms or to conventional wisdom (however defined). Third, reusability necessitates not only additional policies, but also mechanisms for enforcement and compliance (Zimmerman, 2008). Fourth, aligning terms of use or agreements can prove a stumbling block (Smith, 2014). Overall, sharing will likely thrive in communities that realize direct benefits or in situations in which infrastructure is in place (Jones, 2012).

Most important, current eclectic sharing practices contravene best scholarly practices (Nicholson and Bennett, 2011). Wallis *et al.* (2013) report:

The originating investigator bears the cost of data preparation. Other entities such as data repositories, universities, libraries, and funding agencies are likely to bear the cost of curating those data for sustainable access. Unknown – and often nonexistent – reusers reap the benefits. This equation is not viable in economic or social terms (p. 15).

Exacerbating the economic challenges of reuse, the amount of data reuse overall is unknown; in the majority of fields, in fact, demand for reusable data seems close to nil (Borgman, 2015).

## Researchers and researcher behaviors

### *Disciplinary*

“Within or across disciplines,” states Pryor (2012), “members of that workforce will over time combine, disperse and recombine with seeming fluidity; the research undertaken will rarely follow an exclusive and linear path and as a community they will exhibit changing patterns of allegiance and interests” (p. 9). As a result, curation services must embrace a range of interdisciplinary, disciplinary, subdisciplinary and domain practices (Akers and Doty, 2013; Cragin *et al.*, 2010; Karasti *et al.*, 2006; Lage *et al.*, 2011; Lyon, 2007; Lyon and Brenner, 2015; Lyon *et al.*, 2009; Myers *et al.*, 2005; Palmer and Cragin, 2008; Palmer *et al.*, 2009; Pryor, 2009; Shankar, 2007).

It appears that data are not – nor perhaps soon will be – fully translatable among fields (Borgman, 2007). For example, the disciplinary culture of agronomy at one university impeded sharing (Carlson and Stowell-Bracke, 2013). A smattering of studies aside, the ways in which data travel among scholarly fields remains understudied (Edwards *et al.*, 2011).

### *Communities of practice and collaboration*

A community of practice “denotes the level of the social world at which a particular practice is common and coordinated, at which generic understandings are created and shared, and negotiation is conducted” (Davenport and Hall, 2002, p. 172). Communities of practice reconcile an organization's traditional perspective and its evolving plans and priorities (Brown and Duguid, 1991).

Communities of practice revolve around problem solving, knowledge mapping, fulfilling information requests, reusing assets, coordination, documentation, and discussions (Wenger, 2006). They also depend upon collaboration (Wood and Gray, 1991; Yarmey and Baker, 2013). Ideally, collaborations rest upon openness, patience, and tolerance of alternative research philosophies (Gunawardena *et al.*, 2010). Perhaps most important, communities of practice and the collaborative relationships they nourish ground new forms of scholarly communication.

## Scholarly communication

### *Cyberscholarship*

High performance computing, visualization, databases and data sets, and networking enable a new form of scholarly communication: cyberscholarship (Lynch, 2014a). Cyberscholarship rests upon collaborative resource discovery, discussion, and analysis of texts, images (both moving and static), sound recordings, and maps and geographical information systems (Green and Roy, 2008). It allows the pursuit of new forms of research, developing tools for collection-building and analysis, and creating new intellectual products (American Council of Learned Societies, 2006; Arms, 2008). Waters (2007) comments:

Although the systematic exploration of large quantities of information is not a new scholarly practice, what does seem new is the formalization of the very traditional interpretive activities of data-mining, pattern-matching, and simulation in powerful algorithms that represent large and complex sets of data in terms of multiple features and variables that can be analyzed, tested, replicated, and changed at the scale and speed afforded by advanced computation (p. 8).

The development of infrastructure for cyberscholarship demands national, international, and interdisciplinary coordination (Arms and Larsen, 2007). For stakeholders, it requires domain and computational knowledge and a lingua franca as well (Bowker and Star, 2009). Last, cyberscholarship requires strategies to deal with publication and citation as well as with copyright and legal issues. Lynch (2014a) concludes, “Changes in the *practice* of scholarship need to go hand-in-hand with changes in the *communication* and *documentation* of scholarship” (p. 15). Data publication and citation represents a crucial aspect of such communication and documentation.

### *Data publication and citation*

The contention that data sets should be part of the scholarly record and treated as first-class research products continues to attract acolytes (Kratz and Strasser, 2014). Parsons and Fox (2013) characterize data publication as “a metaphor of choice to describe the desired, rigorous, data stewardship approach that creates and curates data as first class objects” (p. WDS32). Consensus obtains: data should be available publicly and in perpetuity; they should be sufficiently documented and validated; and they should be citable (Kratz and Strasser, 2014).

Citations once bolstered scholarly arguments; now they are implicated, too, in provenance, discovery, quality, and attribution (Brase *et al.*, 2014). Citations allow identification, retrieval, and attribution of data; they incentivize sharing and reuse, reputedly promoting scholarly productivity (Mooney and Newton, 2012).

Data citation remains exceptional, however: inertia hampers many research communities (Mayernik, 2012). Researchers may not know that they should cite data. Even should they know, they might not know how to cite them; journals seldom provide instructions. One content analysis found data citation practices desultory (Mooney and Newton, 2012).

Scholars argue that data citations should provide a unique and persistent identifier; receive the same credit as traditional citations; give intellectual and legal attribution; enable access to data, metadata, and documentation; accommodate provenance, standards, and domain practices; and be usable by tools as well as persons (Ball and Duke, 2012; Brase *et al.*, 2014). Outstanding challenges include granularity, microattribution, identifiers for contributions, and citation placement (Ball and Duke, 2012).



*Rights*

Like data publication and citation, copyright and intellectual property are bedeviling issues in digital curation. Common classes of intellectual property include copyright, patents, trademarks, and designs (Corti *et al.*, 2014). For example, copyright law seems both “ubiquitous and yet not at all intuitive” (Levine, 2014, p. 140). But copyright may be less knotty than other intellectual property issues. Smith (2014) sums up, “Copyright law, while complex and nuanced, is largely harmonized ... worldwide, unlike other types of intellectual property law (e.g. sui generis database rights or patent rights)” (p. 46).

Legal issues interpenetrate digital curation work: technological change always outstrips legal change and countries’ copyright laws as well as the disciplinary practices of their researchers vary (Levine, 2014). Options for protecting or sharing data include contracts, public licenses, and waivers (Smith, 2014; Strassser, 2015). Meanwhile “fair use” permits scholars to employ a portion of a copyrighted work for review, critique, or scholarship, for pedagogy, and for archiving (Corti *et al.*, 2014). Fundamental impediments may be more nebulous than copyright, however, namely those surrounding credit and attribution and promotion and tenure (Levine, 2014).

As issues related to copyright and intellectual property, data publication and citation, and cyberscholarship suggest, scholarly communication’s “evolution, revolution, or crisis” continues (Borgman, 2007, p. 9). Infrastructure looms large in this regard.

**Infrastructure**

As Poole (2015) suggests, four types of institutions show much potential with respect to digital curation: archives, centers, libraries, and IRs. Yet many researchers remain underinformed about the strengths and weaknesses of each one (Steinhart *et al.*, 2012).

*Archives*

Archivists’ areas of expertise include ownership, donor relations, policy development, cyberinfrastructure, repository management, intellectual property, selection and appraisal, creation, use, and reuse contexts, authenticity, trustworthiness, provenance, access and use restrictions, permanence, and metadata (Akmon *et al.*, 2011; Dooley, 2015; Gilliland, 2014; Prom, 2011). No wonder Wallis *et al.* (2008) stump for implementing archival practices as early as possible in the digital curation lifecycle.

But archivists’ involvement in digital curation remains exceptional – even as archival principles increasingly permeate sundry academic fields (Dooley, 2015; Manoff, 2004, 2010; Poole, 2015). As Redwine *et al.* (2013) point out, archivists often acquire digital materials incidentally. What is more, archivists face tremendous backlogs of paper materials (Prom, 2011). More than a third (34 percent) of repositories in one study reported that more than half of their holdings remained unprocessed (Greene and Meissner, 2005). Further, archivists may see data as peripheral to historical research and may prefer to deposit them in discipline-specific repositories (Akmon *et al.*, 2011).

Archivists can serve as digital curation participants or consultants or both. A archive’s mission statement and its collection development policy should steer its involvement (Noonan and Chute, 2014). A recent survey of college and university archivists found that nearly half of respondents institution collected institutional or research data. Nonetheless, size mattered: the largest institutions demonstrated the most archivist involvement. Most important, the vast majority of participants (86 percent) believed archivists should be involved with digital curation on some level, but

only 54 percent felt capable of fulfilling their perceived roles (Noonan and Chute, 2014). Despite this discrepancy, archivists are increasingly well-placed to address pressing digital curation needs (Prom, 2011).

### *Centers*

Centers such as the Digital Curation Centre can create or store data, can enable discovery, can promote scholarship, pedagogy, and collaboration, can improve research efficiency and thus increase return on investment, and can raise awareness of data's importance in and across various disciplines and outside the academy (Beagrie, 2004; Collins, 2012; Donnelly, 2013; Hockx-Yu, 2007; Pryor, 2013; Rusbridge, *et al.*, 2005; Zorich, 2009). But still lagging is the ability to embed these centers' expertise institutionally (Lyon, 2007). Indeed, many centers struggle to sustain themselves (Collins, 2012).

### *Libraries*

Campus libraries have a pivotal role to play in supporting and educating researchers (Akers and Green, 2014; Carlson, 2012, 2013; Corral, 2012; Corral *et al.*, 2013; Eaker, 2014; Erdmann, 2013; Fox, 2013; Giarlo, 2013; Gold, 2010; Haendel *et al.*, 2012; Hey and Hey, 2006; Lage *et al.*, 2011; Latham and Poe, 2012; Lyon, 2012; McLure *et al.*, 2014; Mitchell, 2013; Muilenberg *et al.*, 2014; Nicholson and Bennett, 2011; Ovadia, 2013; Repository Task Force, 2009; Scaramozzino *et al.*, 2012; Shaffer, 2013; Starr *et al.*, 2012; Tenopir *et al.*, 2012; Walters, 2009; Walters and Skinner, 2011; Wang, 2013; Westra, 2014; Witt, 2008). As this outpouring of literature implies, "The most popular buzzword in library land these days is *curation*" (Abrams, 2014, p. 25).

Librarians understand metadata, information literacy, scholarly communication, open access, and both collection and repository management (Corral *et al.*, 2013). Libraries and librarians can increase awareness of digital curation's importance, can provide preservation services, and can cultivate new professional practices (Swan and Brown, 2008). Localized studies can help determine the appropriate mix of skills they need (McLure *et al.*, 2014). Perhaps most important, digital curation facilitates librarians' further embedding themselves in research processes (Nicholls *et al.*, 2014). Partnering with disciplinary data repositories such as Dryad is one strategy for doing so (Akers and Green, 2014).

As yet, though, the commitment of libraries to digital curation is inconsistent (Council on Library and Information Resources, 2013). Librarians who work with data often do so more or less in isolation and determining data's place in a library's prosaic operations is hardly intuitive (Steinhart, 2014).

Complicating matters, researchers lament their inability to create sharable metadata and to manage their data efficiently, but show little awareness that libraries and librarians can help (Kroll and Forsman, 2010). Similarly, researchers may not think of librarians as true collaborators (Wright *et al.*, 2014). But DMPs may suffer as a consequence of researchers failing to consult librarians (Hswe and Holt, 2011). As with archivists, librarians' potential with respect to digital curation has only just been tapped (Carlson *et al.*, 2013; Fox, 2013).

*Digital curation profiles (DCPs)*. Developed at Purdue University, DCPs constitute a key resource for librarians and researchers (Witt *et al.*, 2009). DCPs provide information about data creation or use or both, about data management, and about the curation needs of the researcher(s). Comprehensible to non-specialists, they enable comparison across domains and disciplines (Carlson, 2013).

A recent study of DCP workshops determined that 237 of 259 participants (all librarians), had data responsibilities. Participants encountered familiar obstacles, though, primarily organizational support, time, staff, and resources. Although workshops boosted attendees' confidence levels, their levels of engagement changed little (Carlson, 2013).

The DCP initiative also developed a DCPs Toolkit. After users critiqued the DCP Toolkit, Purdue personnel set to developing a new version (Brandt and Kim, 2014; Zhang *et al.*, 2015). Other institutions might profit from Purdue's example.

*DataNet.* Like DCPs building on libraries' potential for digital curation work, the National Science Foundation's 2008 DataNet initiative funded Data Conservancy and DataONE (Lee *et al.*, 2009; Sandusky *et al.*, 2009)[11]. First, Data Conservancy (though defunded in 2013) established a template for developing a distributed human and technical infrastructure (Treloar *et al.*, 2012)[12]. Second, the DataONE project focusses on the biological, ecological, and environmental sciences lifecycles (Allard, 2014; Preservation and Metadata Working Group, n.d.; Tenopir *et al.*, 2011)[13].

Three DataNet projects (2011-2016) succeeded Data Conservancy and DataONE. First, Sustainable Environment through Actionable Data (SEAD) effects "sophisticated management of heterogeneous data while dramatically lowering the cost and effort required to curate and preserve data for long-term community use[14]." Second, "grounded in federated data grid infrastructure," the DataNet Federation Consortium employs the integrated rule-oriented data system to encourage collaborative research among scientists and engineers[15]. Third, *Terra Populus* provides tools for data integration across the domains of social and environmental science[16]. In hosting DataNet projects and developing DCPs, among other activities, libraries seem poised to play a key role in digital curation work. IRs may prove similarly critical. Here, too, librarians play a central if still unsettled role (Allard *et al.*, 2005).

### *IRs*

Often associated with libraries, IRs may support project conception, proposal development, scheduling, documenting, embargoing, communicating within and among research groups, data exchange, and storage (Kunda and Anderson-Wilk, 2011; Walters, 2014). Not only can they manage scholarship and data, but also software, tools, and code (Cragin *et al.*, 2010; Walters, 2014). In choosing a repository, scholars might consider five questions: are similar datasets already stored in the repository? What are the repository's access and use policies? How long will the repository keep the data? Who manages the repository? What costs are associated with using the repository? (Strasser, 2015).

The Purdue University Research Repository (PURR) and the Merritt Repository hosted by the University of California exemplify optimal repository work. PURR seeks to offer a "'cradle-to-grave' service" that focusses on lifecycle services, consultations, publishing, and data discovery and preservation (Brandt, 2014, p. 333). The Merritt meanwhile offers storage, integrity checking, versioning, a metadata catalog, access control rules, use agreements, preservation services, and asynchronous delivery (Abrams *et al.*, 2014).

As PURR and the Merritt suggest, IRs may bridge personal and university servers and national and international repositories (Akers and Green, 2014). Macdonald and Martinez-Uribe (2010) call for further intra-institutional, inter-institutional, and cross-facility services.

Given the overall scale and complexity of the digital curation mandate, there are surely roles for IRs, libraries, centers, and archives. But institutions and organizations need to reconcile variegated disciplinary, domain, and research community perspectives and practices (Lynch, 2014b).

### Overarching concerns

Overarching concerns in digital curation include standards, governance and policy, planning, risk management, evaluation, and metrics, sustainability, and outreach.

#### *Standards*

Standards “facilitate the exchange and comparability of information and practices” (Yarmey and Baker, 2013, p. 157). Key in rendering local into public knowledge, they shore up authority, authenticity, reliability, and usability (Higgins, 2009). Standards smooth the way for discoverability, accessibility, interoperability, collaboration, and preservation (Higgins, 2009; Lyon, 2007; Zimmerman, 2008).

Yet Lack of coordination hinders the development and adoption of standards and best practices, as do broader political, cultural, scientific, and technical issues (Griffiths, 2009; National Research Council of the National Academies, 2015; Zimmerman, 2008). Many disciplines and domains lack standards and remain ill-informed about standards’ affordances (Dallmeier-Tiessen *et al.*, 2014; Lyon, 2007; Tenopir *et al.*, 2011). Choosing among sometimes competing standards adds yet another wrinkle (Dallmeier-Tiessen *et al.*, 2014). One study discerned that the absence of standards hampered graduate students’ efforts to manage their agronomic data (Carlson and Stowell-Bracke, 2013).

Without incentives, moreover, researchers may well resist (Baker and Yarmey, 2009; Edwards, 2004). Despite the potential cost financially and temporally, Yarmey and Baker (2013) lobby for standards-making as an evolutionary process. Standards consistently justify their costs; they encourage the development of local best practices (Lynch, 2008; Yarmey and Baker, 2013). Still, they may cause as many problems as they solve (Borgman, 2015).

#### *Governance and policy*

“The system of decision rights and responsibilities covering who can take what actions with what data, when, under what circumstances and using what methods,” data governance ensures that data can be trusted and that stakeholders remain accountable (Smith, 2014, pp. 45-46). Governance centers on risk management, legal and policy issues, attribution and citation, archiving and preservation, discovery and provenance, data schema and ontologies, and infrastructure (Smith, 2014).

In digital curation, policy vacuums abound with respect to copyright and other intellectual property issues (Smith, 2014). Researchers negotiate Institutional Review Board and funders’ requirements, professional ethical codes, and cultural differences (Council on Library and Information Resources, 2013). But policies often gloss over or ignore inter-domain differences in research practices (Borgman, 2015). They tend also to emphasize supplying more data, probing neither scholarly motivations nor expected reuse demand nor necessary infrastructural investment (Borgman, 2015). In some cases, policies are at loggerheads.

Two recent studies illuminate the inchoate state of policy development. First, only 9 percent of respondents to the DataRes Project’s survey said their institution had a data management policy. Nearly three-quarters, by contrast (72 percent), said their institutions did not have such a policy and fully 19 percent said they did not know

(Council on Library and Information Resources, 2013). Second, Dietrich *et al.* (2012) found that funders' policies often lacked specific language and emphasized access to more than preservation of data – even though they provided little guidance about enabling access. Policies also neglected by and large to address open access.

Policies should address accountability, legitimacy, best practices, and monitoring and review (Harvey, 2010). They should accommodate various types of institutions and data products (Lyon, 2007). Though institutional policies should align with funding agency mandates, both institutional inertia and the still-ambiguous status of data qua research product undercut such efforts (Council on Library and Information Resources, 2013). Ultimately, judicious planning periodically reviewed can mold governance and policy.

### *Planning and DMPs*

DMP requirements reflect funders' efforts to show salutary economic and societal impact (Lynch, 2014b). Following the lead of the National Institutes of Health (2003), the National Science Foundation mandated that all grant applications beginning in 2011 include a DMP. This represented the first step in ensuring that federally funded research would become available to the general public (Mervis, 2010). In implementing its own DMP requirements, Holdren's memo impelled researchers to make the results of federally funded scientific research available to maximize the impact of the federal government's investment and to ensure the accountability of those funded.

Four studies indicate the ambiguous status of DMPs in researchers' work. First, the DataRes project's survey found that the vast majority (87 percent) of respondents agreed or strongly agreed that a DMP was important (Council on Library and Information Resources, 2013). But Steinhart *et al.*'s (2012) study at Cornell University found great uncertainty among Principal Investigators about National Science Foundation requirements. Nearly two-thirds (62 percent) of respondents wanted help writing DMPs (only 13 percent expressed no interest). A third study found many researchers at Colorado State University ignorant of what a DMP even was. Researchers also held varied perspectives as to what a DMP mandated in practice (McLure *et al.*, 2014). A fourth study (182 DMPs authored by University of Minnesota scholars who had received NSF grants) determined that a wide variety of often inconsistent practices existed in terms of how scholars shared their data, where they shared it, possible reusers, and mechanisms to allow long-term reuse (Bishoff and Johnston, 2015). Librarians can conduct workshops disseminate best practices, foster communication, and promote tools (Nicholls *et al.*, 2014).

Despite researchers' professed commitment, planning may be postponed or even dismissed (Donnelly, 2012). Many researchers neglect to develop DMPs unless prodded by funding agencies and funders themselves struggle to track compliance (Hswe and Holt, 2011). Barring vehicles for enforcement, audits may prevail – hardly an ideal option (Lynch, 2014b).

Key considerations in crafting a DMP include the data to be generated and shared, funder, publisher, and internal requirements, disciplinary and domain standards and best practices, relevant policies, documentation, deposit location and provisions for backup, legal and ethical obligations, goals for sharing, roles and responsibilities of stakeholders, and resources needed (Corti *et al.*, 2014; Donnelly, 2012; Ray, 2014; Strasser, 2015; Van den Eynden, *et al.*, 2010). A DMP should address stakeholders ranging from researchers, librarians, and center or repository managers, to support

staff such as grants officers, specialists such as archivists, and technical and laboratory staff (Donnelly, 2012)[17]. Finally, a DMP must remain a document amenable to periodic review and refinement (Strasser, 2015).

*Planning tools.* Researchers' current practices appear fragmented largely because funding agencies propagate broad requirements and provide few resources (Sallans and Lake, 2014). A vicious circle may prevail: most disciplines lack standardized procedures for managing data and most researchers have yet to skill up (Sallans and Lake, 2014). Tools such as DMPonline and DMPTool can help.

Building upon librarians' expertise, DMPonline and DMPTool both provide structured environments for data management planning and link to funder requirements (Sallans and Donnelly, 2012). "While the development paths of the tools have diverged," Sallans and Donnelly (2012) conclude, "both groups retain a broader vision of a joined-up tool [...] that serves as a coordinating hub for the management of data across many disciplines, many funding agencies, many institutions and many countries, with shared good practice as a common goal" (p. 128).

Recently the six original partner institutions involved in DMPTool developed DMPTool 2. Goals for DMPTool 2 included settling upon best practices, allowing local adaptation, cultivating an open source community, and promoting the sharing of institutional resources (Strasser *et al.*, 2014). Similarly, DMPonline's fourth version rendered the tool even more user-friendly (Getler *et al.*, 2014).

#### *Risk management, evaluation, and metrics*

In digital curation, risk management constitutes a dynamic endeavor that seeks to limit and control risks to specific activities and assets (Barateiro, *et al.*, 2010). Risk management procedures include discerning potential risks, assessing the degrees of risk, gathering information, settling upon controls, reporting results, and developing action plans (Price and Smith, 2000).

Minimizing risks to data involves providing for persistent access, multiple copies, periodic backup, a delimited number of file formats, secure storage, disaster recovery, and community monitoring (Harvey, 2010). Further, in choosing a backup strategy stakeholders should assess local conditions, the value of the data, the systems that store them, and acceptable levels of risk (Van den Eynden *et al.*, 2010).

Risk management funnels into evaluation. Evaluation stimulates a culture of quality and responsibility (Lakos and Phipps, 2004). Audits allow stakeholders to identify previously unnoticed risks and to refine their definitions of those designated communities whom they serve (Reilly and Waltz, 2014). Four tools lend themselves to this process.

First, the *Trustworthy Repositories Audit & Certification: Criteria and Checklist* (TRAC) helps repositories transparently tackle risks. Although it does not stipulate standards and specifications, on one hand, it insists that the designated community's priorities are non-negotiable, on the other (Center for Research Libraries, 2007; Reilly and Waltz, 2014). Second, the Planning Tool for Trusted Electronic Repositories helps repositories establish trusted status by assisting in the development of performance objectives (DigitalPreservationEurope, 2008). Third, the Digital Repository Audit Method Based on Risk Assessment allows repositories to self-assess and thus to quantify their risks (McHugh *et al.*, 2008). Finally, the Consultative Committee for Space Data Systems (CCSDS)'s 2011 Audit and Certification of Trustworthy Digital Repositories premises continuous improvement (It became ISO 16363 in 2012).

Such tools cannot inoculate an organization against risk, however; measurement is crucial. Measurement facilitates learning, decision making, and awareness-raising (Whyte *et al.*, 2014). It should focus both on developing institutions' research support services and infrastructure and on ultimately describing their economic impact (Whyte *et al.*, 2014). As benefits can resist measurement, a dynamic mix of strategies that incorporates qualitative as well as quantitative analysis is necessary (Fry *et al.*, 2008). Metrics can be leveraged to provide for sustainability.

### *Sustainability*

Inadequate financial resources loom large (National Research Council of the National Academies, 2015). Counterintuitively, quantifying benefits may be more difficult than quantifying costs (Fry *et al.*, 2008). Though Lavoie (2012) argues that sustainability is inextricable from economics, the Repository Task Force (2009) insists that sustainability involves organizational commitment and collaboration. At any rate, optimal sustainability plans feature value propositions, varied revenue streams, and commitments to accountability (Maron *et al.*, 2009).

Projects may assume two approaches to costing. First, the project can price all activities and resources for its entire lifecycle. This allows the calculation of the total cost of data generation, sharing, and preservation. Second, the project can determine the additional expenses that will be incurred to allow the data to be sharable (Van den Eynden *et al.*, 2010).

The LIFE Project, for example, improved planning, comparison, and evaluation of digital lifecycles (LIFE Project Team, 2008). The Keeping Research Data Safe (KRDS) project constructed an activity model and a TRAC-based resource template (Beagrie *et al.*, 2008). KRDS 2 reviewed and extended KRDS's activity model and crafted a benefits framework based on case studies (Beagrie *et al.*, 2010). Finally, the European Union's 4C: Collaborating to Clarify the Cost of Curation project developed the Digital Curation Sustainability Model and the Curation Cost Concept Model and built a Curation Costs Exchange (Kilbride and Norris, 2014; Middleton, 2014).

Ultimately, the sustainability of digital curation work should focus on making decisions that are optimal for current needs and yet will not preclude future stakeholders from changing course (Kunda and Anderson-Wilk, 2011). A dearth of funding imperils data and saps the development of a competent workforce to ensure data's proper curation (National Research Council of the National Academies, 2015). Lyon (2007) insists, "We need to construct new economic models to provide a robust foundation to funding plans, models which link research strategy and program development, to operational support and infrastructure provision." In addition to financial concerns, sustainability incorporates raising awareness and outreach.

### *Outreach*

Both cognizance and comprehension of digital curation best practices remains low (Lyon, 2007). Researchers tend to satisfice, as many feel that funding agencies, publishers, and professional societies neglect to extend sufficient support (Carlson *et al.*, 2013; Kroll and Forsman, 2010). Researchers who participated in a study at Emory University emphasized their interest in attending digital curation workshops and in receiving help preparing their DMPs (Akers and Doty, 2013). More generally, digital curation professionals can promote their web presence and leverage social media, can cultivate intra-institutional collaborations and work with the Office of Research to raise awareness, can serve as liaisons for researcher inquiries, can present at departmental meetings, host information sessions or events, and disseminate

promotional materials, can encourage data deposit, and can meet with researchers new to the institution (Erway *et al.*, 2016).

Only appropriate incentives can propel cultural change in research communities (Jones, 2012; Mooney and Newton, 2012; Whitlock, 2011). “Researchers,” conclude Patrick and Wilson (2013) astutely, “are not rebellious schoolchildren who need to be bullied into working harder; they are generally highly motivated and highly skilled individuals who take a great deal of pride in what they do, and thus are more likely to embrace digital curation as a worthy goal if persuaded of its merits.” Outreach and engagement with actual user behaviors is imperative to improve services (Reilly and Waltz, 2014). Outreach, after all, necessitates little if any further financial investment – merely a receptive audience (Shorish, 2012).

### Directions for further research

“We are at the early stages of a genuine systemic and systematic response to the data stewardship challenges framed by the emergence of e-research, and to seizing the opportunities promised by more effective, broadscale sharing and reuse,” Lynch (2014b) contends (p. 406). In addressing this transition, scholars might look to six questions. First, at what point(s) of the lifecycle can digital curators’ input and support be most valuable? Second, what can further examination of researcher practices suggest about common digital curation needs across communities of practice, domains, subdisciplines, disciplines, institutions, and nations? Third, how can archivists best expand their incipient role in digital curation work? Fourth, what strategies can be used to secure funding beyond one-shot grants and what strategies can be employed to secure funding for ongoing as well as for new digital curation initiatives? Fifth, what outreach and awareness-raising strategies have proven most successful? Finally, what are the most appropriate metrics for measuring success in digital curation work?

“A revolution in digital information is occurring across all realms of human endeavor,” notes the National Research Council of the National Academies (2015, p. 7). Digital curation can help put this revolution to work.

### Notes

1. [www-01.ibm.com/software/data/bigdata/what-is-big-data.html](http://www-01.ibm.com/software/data/bigdata/what-is-big-data.html)
2. “A person, company, etc., with a concern or (esp. financial) interest in ensuring the success of an organization, business, system, etc.” (OED).
3. Digital Curation Center Glossary: [www.dcc.ac.uk/digital-curation/what-digital-curation](http://www.dcc.ac.uk/digital-curation/what-digital-curation)
4. [www.ukoln.ac.uk/projects/I2S2/documents/I2S2-ResearchActivityLifecycleModel-110407.pdf](http://www.ukoln.ac.uk/projects/I2S2/documents/I2S2-ResearchActivityLifecycleModel-110407.pdf)
5. [www.ddialliance.org/Specification/DDI-Lifecycle/3.2/](http://www.ddialliance.org/Specification/DDI-Lifecycle/3.2/)
6. [www.ijdc.net/index.php/ijdc/article/view/133](http://www.ijdc.net/index.php/ijdc/article/view/133)
7. [www.cell.com/trends/ecology-evolution/pdf/S0169-5347%2811%2900339-9.pdf](http://www.cell.com/trends/ecology-evolution/pdf/S0169-5347%2811%2900339-9.pdf)
8. [www.2.le.ac.uk/services/research-data/images/new\\_institution.PNG/view](http://www.2.le.ac.uk/services/research-data/images/new_institution.PNG/view)
9. [www.asis.org/asist2011/proceedings/submissions/36\\_FINAL\\_SUBMISSION.pdf](http://www.asis.org/asist2011/proceedings/submissions/36_FINAL_SUBMISSION.pdf)
10. [www.data-archive.ac.uk/create-manage/life-cycle](http://www.data-archive.ac.uk/create-manage/life-cycle)
11. [www.nsf.gov/pubs/2007/nsf07601/nsf07601.htm](http://www.nsf.gov/pubs/2007/nsf07601/nsf07601.htm)
12. <https://dataconservancy.org/>



13. [www.dataone.org/what-dataone](http://www.dataone.org/what-dataone)
14. <http://sead-data.net/>
15. <http://datafed.org/about/>
16. [www.terrapop.org/](http://www.terrapop.org/)
17. Useful guides include the DCC's DMPT Online: Data Management Planning Tool (<http://dmponline.dcc.ac.uk/>), Martin Donnelly and Sarah Jones. "Checklist for a Data Management Plan" ([www.dcc.ac.uk/sites/default/files/documents/data-forum/documents/docs/DCC\\_Checklist\\_DMP\\_v3.pdf](http://www.dcc.ac.uk/sites/default/files/documents/data-forum/documents/docs/DCC_Checklist_DMP_v3.pdf)), the ICPSR's "Guidelines for Effective Data Management Plans" ([www.icpsr.umich.edu/icpsrweb/ICPSR/dmp/index.jsp](http://www.icpsr.umich.edu/icpsrweb/ICPSR/dmp/index.jsp)), the California Digital Library's "DMPTool: Guidance and Resources for Your Data Management Plan," and the UK Data Archives's "Managing and Sharing Data" ([www.data-archive.ac.uk/media/2894/managingsharing.pdf](http://www.data-archive.ac.uk/media/2894/managingsharing.pdf)).

### References

- Abbott, D. (2015), "Digital curation and doctoral research: current practice", *International Journal of Digital Curation*, Vol. 10 No. 1, pp. 1-17.
- Abrams, S. (2014), "Curation: buzzword or what?", *Information Outlook*, Vol. 18 No. 5, pp. 25-27.
- Abrams, S., Cruse, P., Strasser, C., Willet, P., Boushey, G., Kochi, J., Laurance, M. and Rizk-Jackson, A. (2014), "DataShare: empowering researcher data curation", *International Journal of Digital Curation*, Vol. 9 No. 1, pp. 110-118.
- Akers, K. and Doty, J. (2013), "Disciplinary differences in faculty research data management practices and perspectives", *International Journal of Digital Curation*, Vol. 8 No. 2, pp. 5-26.
- Akers, K. and Green, J.A. (2014), "Towards a symbiotic relationship between academic libraries and disciplinary data repositories: a Dryad and University of Michigan case study", *International Journal of Digital Curation*, Vol. 9 No. 1, pp. 119-131.
- Akmon, D., Zimmerman, A., Daniels, M. and Hedstrom, M. (2011), "The application of archival concepts to a data-intensive environment: working with scientists to understand data management and preservation needs", *Archival Science*, Vol. 11 No. 3, pp. 329-348.
- Allard, S. (2014), "Evaluating a complex project: DataONE", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 255-274.
- Allard, S., Mack, T.R. and Feltner-Reichert, M. (2005), "The librarian's role in institutional repositories: a content analysis of the literature", *Reference Services Review*, Vol. 33 No. 3, pp. 325-336.
- American Council of Learned Societies (2006), *Our Cultural Commonwealth: The Report of the American Council of Learned Societies Commission on Cyberinfrastructure for the Humanities and Social Sciences*, American Council of Learned Societies and Andrew W. Mellon Foundation, New York, NY.
- Anderson, J. and Rainie, L. (2012), *Big Data: Experts Say New Forms of Information Analysis will Help People be More Nimble and Adaptive, but Worry over Humans' Capacity Tounderstand and Use these New Tools Well*, Pew Research Center, Washington, DC.
- Arms, W.Y. (2008), "Cyberscholarship: high performance computing meets digital libraries", *Journal of Electronic Publishing*, Vol. 11 No. 1, available at: <http://quod.lib.umich.edu/j/jep/3336451.0011.103?view=text;rgn=main> (accessed August 20, 2016).
- Arms, W.Y. and Larsen, R.L. (2007), *The Future of Scholarly Communication: Building the Infrastructure for Cyberscholarship*, National Science Foundation and Joint Information Systems Committee, Phoenix, AZ.

- Baker, K.S. and Yarmey, L. (2009), "Data stewardship: environmental data curation and a web-of-repositories", *International Journal of Digital Curation*, Vol. 4 No. 2, pp. 1-16.
- Ball, A. (2012), *Review of Data Management Lifecycle Models*, University of Bath, Bath.
- Ball, A. and Duke, M. (2012), *How to Cite Datasets and Link to Publications*, Digital Curation Center, Edinburgh.
- Barateiro, J., Antunes, G., Freitas, F. and Borbinha, J. (2010), "Designing digital preservation solutions: a risk management-based approach", *International Journal of Digital Curation*, Vol. 5 No. 1, pp. 4-17.
- Beagrie, N. (2004), "The digital curation centre", *Learned Publishing*, Vol. 17 No. 1, pp. 7-9.
- Beagrie, N., Chruszcz, J. and Lavoie, B. (2008), "Keeping research data safe: a cost model", Joint Information Systems Committee (JISC), Salisbury.
- Beagrie, N., Lavoie, B. and Woollard, M. (2010), "Keeping research data safe 2", Joint Information Systems Committee (JISC), Salisbury.
- Birnholtz, J.P. and Bietz, M.J. (2003), "Data at work: supporting sharing in science and engineering", *ACM, Sanibel Island, FL*, pp. 339-348.
- Bishoff, C. and Johnston, L. (2015), "Approaches to data sharing: an analysis of NSF data management plans from a large research university", *Journal of Librarianship and Scholarly Communication*, Vol. 3 No. 2, pp. 1-27.
- Bollacker, K. (2010), "Avoiding a digital dark age", *American Scientist*, Vol. 98 No. 2, pp. 106-110.
- Borgman, C. (2007), *Scholarship in the Digital Age*, MIT Press, Cambridge.
- Borgman, C. (2012), "The conundrum of sharing research data", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 6, pp. 1059-1078.
- Borgman, C. (2015), *Big Data, Little Data, No Data*, MIT Press, Cambridge, MA.
- Borgman, C.L., Wallis, J.C. and Mayernik, M. (2012), "Who's got the data? Interdependencies in science and technology collaborations", *Computer Supported Cooperative Work*, Vol. 21 No. 6, pp. 485-523.
- Bowker, G.C. and Star, S.L. (2009), "Cyberscholarship; or, 'A Rose is a Rose is a...'", *EDUCAUSE Review*, Vol. 44 No. 3, pp. 6-7.
- Brandt, D.S. (2014), "Purdue University Research Repository", in Ray, J. (Ed.), *Research Data Management*, Purdue University Press, West Lafayette, IN, pp. 325-345.
- Brandt, D.S. and Kim, E. (2014), "Data curation profiles as a means to explore managing, sharing, disseminating or preserving digital outcomes", *International Journal of Performance Arts and Digital Media*, Vol. 10 No. 1, pp. 21-34.
- Brase, J., Socha, Y., Callaghan, S., Borgman, C.L., Uhlir, P.F. and Carroll, B. (2014), "Data citation: principles and practice", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 167-186.
- Brown, J.S. and Duguid, P. (1991), "Organizational knowledge and communities-of-practice: toward a unified view of working, learning, and innovation", *Organization Science*, Vol. 2 No. 1, pp. 40-57.
- Carlson, J. (2012), "Demystifying the data interview", *Reference Services Review*, Vol. 40 No. 1, pp. 7-23.
- Carlson, J. (2013), "Opportunities and barriers for librarians in exploring data: observations from the data curation profile workshops", *Journal of eScience Librarianship*, Vol. 2 No. 2, pp. 17-33.
- Carlson, J. (2014), "The use of lifecycle models", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 63-86.

- Carlson, J. and Stowell-Bracke, M. (2013), "Data management and sharing from the perspective of graduate students: an examination of the culture and practice at the water quality field station", *Portal: Libraries and the Academy*, Vol. 13 No. 4, pp. 343-361.
- Carlson, J., Fosmire, M., Miller, C. and Nelson, M.S. (2011), "Determining data information literacy needs: a study of students and research faculty", *Portal: Libraries and the Academy*, Vol. 11 No. 2, pp. 629-657.
- Carlson, J., Johnston, L., Westra, B. and Nichols, M. (2013), "Developing an approach for data management education: a report from the data information literacy project", *International Journal of Digital Curation*, Vol. 8 No. 1, pp. 204-217.
- Carlson, S. and Anderson, B. (2007), "What are data? The many kinds of data and their implications for data re-use", *Journal of Computer-Mediated Communication*, Vol. 12 No. 2, pp. 301-317.
- Center for Research Libraries (2007), *Trustworthy Repositories Audit & Certification: Criteria and Checklist*, OCLC and CRL, Dublin.
- Collins, E. (2012), "The national data centres", in Pryor, G. (Ed.), *Research Data Management*, Facet, London, pp. 151-172.
- Constantopoulos, P. and Dallas, C. (2010), "Aspects of a digital curation agenda for digital heritage", IEEE, Athens.
- Corrall, S. (2012), "Roles and responsibilities: libraries, librarians and data", in Pryor, G. (Ed.), *Managing Research Data*, Facet, London, pp. 105-133.
- Corrall, S., Kennan, M.A. and Afzal, W. (2013), "Bibliometrics and research data management services: emerging trends in library support for research", *Library Trends*, Vol. 61 No. 3, pp. 636-674.
- Corti, L., Van Den Eynden, V., Bishop, L. and Woolard, M. (2014), *Managing and Sharing Research Data: A Guide to Good Practice*, Sage, Los Angeles, CA.
- Council on Library and Information Resources (2013), *Research Data Management: Principles, Practices, and Prospects*, Council on Library and Information Resources, Washington, DC.
- Cragin, M., Palmer, C.L., Carlson, J.R. and Witt, M. (2010), "Data sharing, small science and institutional repositories", *Transactions of the Royal Society*, Vol. 368 No. 1926, pp. 4023-4038.
- Dallas, C. (2007), *An Agency-Oriented Approach to Digital Curation Theory and Practice*, Archives and Museum Informatics, Toronto.
- Dallmeier-Tiessen, S., Darby, R., Gitmans, K., Lambert, S., Matthews, B., Mele, S., Suhonen, J. and Wilson, M. (2014), "Enabling sharing and reuse of scientific data", *New Review of Information Networking*, Vol. 19 No. 1, pp. 16-43.
- Davenport, E. and Hall, H. (2002), "Organizational knowledge and communities of practice", *Annual Review of Information Science and Technology*, Vol. 36 No. 1, pp. 171-227.
- Dietrich, D., Adams, T.M.A. and Steinhart, G. (2012), "De-mystifying the data management requirements of funders", *Issues in Science and Technology Librarianship*, Vol. 70 No. 1, available at: [www.istl.org/12-summer/refereed1.html](http://www.istl.org/12-summer/refereed1.html) (accessed August 20, 2016).
- DigitalPreservationEurope (2008), "Repository planning checklist and guidance DPE-D3.2", HATII, Glasgow.
- Donnelly, M. (2012), "Data management plans and planning", in Pryor, G. (Ed.), *Managing Research Data*, Facet, London, pp. 83-103.
- Donnelly, M. (2013), "The DCC's institutional engagements: raising research data management capacity in UK higher education", *Bulletin of the Association for Information Science and Technology*, Vol. 39 No. 6, pp. 37-40.

- Dooley, J. (2015), *The Archival Advantage: Integrating Archival Expertise into Management of Born-Digital Library Materials*, OCLC Research, Dublin, OH.
- Duerr, R., Parsons, M.A., Marquis, M., Dichtl, R. and Mullins, T. (2004), "Challenges in long-term data stewardship", *IEEE*, pp. 101-121.
- Eaker, C. (2014), "Planning data management education initiatives: process, feedback, and future directions", *Journal of eScience Librarianship*, Vol. 3 No. 1, pp. 3-14.
- Education Advisory Board (2014), *Market Demand for Digital Curation Master's Degree Programs*, Education Advisory Board, Washington, DC.
- Edwards, P.N. (2004), "'A vast machine': standards as social technology", *Science*, Vol. 304 No. 5672, pp. 827-828.
- Edwards, P.N. *et al.* (2011), "Science friction: data, metadata, and collaboration", *Social Studies of Science*, Vol. 41 No. 5, pp. 667-690.
- Erdmann, C. (2013), "Teaching librarians to be data scientists", *Information Outlook*, Vol. 18 No. 3, pp. 21-24.
- Erway, R., Horton, L., Nurnberger, A., Osuji, R. and Rushing, A. (2016), *Building Blocks: Laying the Foundation for a Research Data Management Program*, OCLC Research, Dublin, OH.
- Faniel, I. and Jacobsen, T. (2010), "Reusing scientific data: how earthquake engineering researchers assess the reusability of colleagues' data", *Computer Supported Cooperative Work*, Vol. 19 No. 3, pp. 355-375.
- Faniel, I.M. and Zimmerman, A. (2011), "Beyond the data deluge: a research agenda for large-scale data sharing and reuse", *International Journal of Digital Curation*, Vol. 6 No. 1, pp. 58-69.
- Fox, R. (2013), "The art and science of data curation", *OCLC Systems and Services*, Vol. 29 No. 4, pp. 195-199.
- Fry, J., Lockyer, S., Oppenheim, C., Houghton, J. and Rasmussen, B. (2008), *Identifying Benefits Arising From the Curation and Open Sharing of Research Data Produced by UK Higher Education and Research Institutes*, Loughborough University/Centre for Centre for Strategic Economic Studies, Loughborough.
- Gantz, J. *et al.* (2008), "The diverse and exploding digital universe: an updated forecast of worldwide information growth through 2011", IDC, Framingham, MA, available at: [www.ifap.ru/library/book268.pdf](http://www.ifap.ru/library/book268.pdf) (accessed August 20, 2016).
- Getler, M., Jones, S., Sisu, D. and Miuller, K. (2014), "DMPonline version 4.0: user-led innovation", *International Journal of Digital Curation*, Vol. 9 No. 1, pp. 193-219.
- Giarlo, M. (2013), "Academic libraries as data quality hubs", *Journal of Librarianship and Scholarly Communication*, Vol. 1 No. 3, pp. 1-10.
- Gilliland, A. (2014), *Conceptualizing 21st Century Archives*, Society of American Archivists, Chicago, IL.
- Gold, A. (2010), *Data Curation and Libraries: Short-Term Developments, Long-Term Prospects*, California Polytechnic State University, San Luis Obispo, CA.
- Greenberg, J. (2005), "Understanding metadata and metadata schemes", *Cataloging and Classification Quarterly*, Vol. 40 Nos 3/4, pp. 17-36.
- Green, D. and Roy, M. (2008), "Things to do while waiting for the future to happen: building cyberinfrastructure for the liberal arts", *EDUCAUSE Review*, Vol. 43 No. 4, pp. 35-48.
- Greene, M.A. and Meissner, D. (2005), "More product, less process: revamping traditional archival processing", *American Archivist*, Vol. 68 No. 2, pp. 208-263.
- Griffiths, A. (2009), "The publication of research data: research attitudes and behavior", *International Journal of Digital Curation*, Vol. 4 No. 1, pp. 46-56.

- Gunawardena, S., Weber, R. and Agosto, D.E. (2010), "Finding that special someone: interdisciplinary collaboration in an academic context", *Journal of Education for Library and Information Science*, Vol. 51 No. 4, pp. 210-221.
- Haendel, M.A., Vasilevsky, N.A. and Wirz, J.A. (2012), "Dealing with data: a case study on information and data management literacy", *PLoS Biology*, Vol. 10 No. 5, pp. 1-4.
- Harvey, R. (2010), *Digital Curation: A How to Do It Manual*, Neal Schuman, New York, NY.
- Heidorn, P.B. (2008), "Shedding light on the dark data in the long tail of science", *Library Trends*, Vol. 57 No. 2, pp. 280-299.
- Heritage Preservation (2005), *A Public at Risk: The Heritage Health Index Report on the State of America's Collections; A Project of Heritage Preservation and the Institute of Museum and Library Services*, Heritage Preservation, Washington, DC.
- Hey, T. and Hey, J. (2006), "e-Science and its implications for the library community", *Library Hi Tech*, Vol. 24 No. 4, pp. 515-528.
- Higgins, S. (2008), "The DCC curation lifecycle model", *International Journal of Digital Curation*, Vol. 3 No. 1, pp. 134-140.
- Higgins, S. (2009), "DCC DIFFUSE standards frameworks: a standards path through the curation lifecycle", *International Journal of Digital Curation*, Vol. 4 No. 2, pp. 60-67.
- Hockx-Yu, H. (2007), "Digital curation centre – phase two", *International Journal of Digital Curation*, Vol. 2 No. 1, pp. 123-127.
- Holdren, J.P. (2013), "Increasing access to the results of federally funded scientific research", available at: [www.whitehouse.gov/sites/default/files/microsites/ostp/ostp\\_public\\_access\\_memo\\_2013.pdf](http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf) (accessed August 15, 2016).
- Hswe, P. and Holt, A. (2011), "Joining in the enterprise of response in the wake of the NSF data management planning requirement", *Research Library Issues*, Vol. 274, pp. 11-16.
- Jahnke, L. and Asher, A. (2012), *The Problem of Data*, Council on Library and Information Resources, Washington, DC.
- Jones, S. (2012), "Research data policies", in Pryor, G. (Ed.), *Research Data Management*, Facet, London, pp. 47-66.
- Karasti, H., Baker, K.S. and Halkola, E. (2006), "Enriching the notion of data curation in E-science: data managing and information infrastructuring in the long term ecological research (LTER) network", *Computer Supported Cooperative Work*, Vol. 15 No. 4, pp. 321-358.
- Kilbride, W. and Norris, S. (2014), "Collaborating to classify the cost of curation", *New Review of Information Networking*, Vol. 19 No. 1, pp. 44-48.
- Kratz, J. and Strasser, C. (2014), "Data publication consensus and controversies", *F1000Research*, Vol. 3 No. 94, available at: <http://f1000research.com/articles/3-94/v1> (accessed August 20, 2016).
- Kroll, S. and Forsman, R. (2010), *A Slice of Research Life: Information Support for Research in the United States*, OCLC Research, Dublin, OH.
- Kunda, S. and Anderson-Wilk, M. (2011), "Community stories and institutional stewardship: digital curation's dual roles of story creation and resource preservation", *Portal: Libraries and the Academy*, Vol. 11 No. 4, pp. 895-914.
- Lage, K., Losoff, B. and Maness, J. (2011), "Receptivity to library involvement in scientific data curation: a case study at the University of Colorado Boulder", *Portal: Libraries and the Academy*, Vol. 11 No. 4, pp. 915-937.
- Lakos, A. and Phipps, S. (2004), "Creating a culture of assessment: a catalyst for organizational change", *Portal: Libraries and the Academy*, Vol. 4 No. 3, pp. 345-361.
- Latham, B. and Poe, J.W. (2012), "The library as partner in university data curation: a case study in collaboration", *Journal of Web Librarianship*, Vol. 6 No. 4, pp. 288-304.

- Lavoie, B.F. (2012), "Sustainable research data", in Pryor, G. (Ed.), *Research Data Management*, Facet, London, pp. 67-82.
- Lee, C.A. (2005), *Defining Digital Preservation Work: A Case Study of the Development of the Reference Model for an Open Archival Information System*, University of Michigan, Ann Arbor, MI.
- Lee, C.A. (2009), "Open archival information system (OAIS) reference model", in Bates, M.J. and Maack, M.N. (Eds), *Encyclopedia of Library and Information Sciences*, CRC Press, Boca Raton, FL, pp. 4020-4030.
- Lee, C.A. and Tibbo, H.R. (2007), "Digital curation and trusted repositories: steps toward success", *Journal of Digital Information*, Vol. 8 No. 2, available at: <https://journals.tdl.org/jodi/index.php/jodi/article/view/229/183> (accessed August 18, 2016).
- Lee, J., Zhang, J., Zimmerman, A. and Lucia, A. (2009), "DataNet: an emerging cyberinfrastructure for sharing, reusing and preserving digital data for scientific discovery and learning", *American Institute of Chemical Engineers Journal*, Vol. 55 No. 11, pp. 2757-2764.
- Levine, M. (2014), "Copyright, open data, and the availability-usability gap: challenges, opportunities, and approaches for libraries", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 129-147.
- LIFE Project Team (2008), *The LIFE2 Final Project Report*, Joint Information Systems Committee, London.
- Lynch, C. (2008), "Big data: how do your data grow?", *Nature*, Vol. 455 No. 7209, pp. 28-29.
- Lynch, C. (2014a), "The 'digital scholarship' disconnect", *EDUCAUSE Review*, pp. 10-15.
- Lynch, C. (2014b), "The next generation of challenges in the curation of scholarly data", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 395-408.
- Lyon, L. (2007), *Dealing with Data: Roles, Rights, Responsibilities, and Relationships*, UKOLN, University of Bath, Bath.
- Lyon, L. (2009), "Open science at web scale: optimizing participation and predictive potential", Joint Information Systems Committee (JISC), Bath.
- Lyon, L. (2012), "The informatics transform: re-engineering libraries for the data decade", *International Journal of Digital Curation*, Vol. 7 No. 1, pp. 126-138.
- Lyon, L. and Brenner, A. (2015), "Bridging the data talent gap: positioning the iSchool as an agent for change", *International Journal of Digital Curation*, Vol. 10 No. 1, pp. 111-122.
- Lyon, L., Rusbridge, C., Neilson, C. and Whyte, A. (2009), *Disciplinary Approaches to Sharing, Curation, Reuse and Preservation*, Digital Curation Centre, Edinburgh.
- McHugh, A., Ross, S., Innocenti, P., Ruusalepp, P. and Hofman, H. (2008), "Bringing self-assessment home: repository profiling and key lines of enquiry within DRAMBORA", *International Journal of Digital Curation*, Vol. 3 No. 2, pp. 130-142.
- McLeod, J., Childs, S. and Lomas, E. (2013), "Research data management", in Pickard, A.J. (Ed.), *Research Methods in Information*, Neal-Schuman, Chicago, IL, pp. 71-86.
- McLure, M., Level, A.V., Cranston, C.L., Oehlerts, B. and Culbertson, M. (2014), "Data curation: a study of researcher practices and needs", *Portal: Libraries and the Academy*, Vol. 14 No. 2, pp. 139-164.
- Macdonald, S. and Martinez-Uribe, L. (2010), "Collaboration to data curation: harnessing institutional expertise", *New Review of Academic Librarianship*, Vol. 16 No. S1, pp. 4-16.
- Manoff, M. (2004), "Theories of the archive from across the disciplines", *Portal: Libraries and the Academy*, Vol. 4 No. 1, pp. 9-25.

- Manoff, M. (2010), "Archive and database as metaphor: theorizing the historical record", *Portal: Libraries and the Academy*, Vol. 10 No. 4, pp. 385-398.
- Maron, N.L., Smith, K.K. and Loy, M. (2009), "Sustaining digital resources: an on-the-ground view of projects today", ITHAKA, New York, NY.
- Mayernik, M. (2012), "Data citation initiatives and issues", *Bulletin of the American Society for Information Science and Technology*, Vol. 38 No. 5, pp. 23-28.
- Mervis, J. (2010), "NSF to ask every grant applicant for data management plan", *ScienceInsider*, May 5.
- Middleton, S. (2014), "Curation cost exchanges: supporting smarter investments in digital curation", *EDUCAUSE Review*, November 10, available at: <http://er.educause.edu/articles/2014/11/curation-costs-exchange-supporting-smarter-investments-in-digital-curation> (accessed August 20, 2016).
- Mitchell, E.T. (2013), "Research support: the new mission for libraries", *Journal of Web Librarianship*, Vol. 7 No. 1, pp. 109-113.
- Mooney, H. and Newton, M.P. (2012), "The anatomy of a data citation: discovery, reuse, and credit", *Journal of Librarianship and Scholarly Communication*, Vol. 1 No. 1, available at: <http://jlscc-pub.org/articles/abstract/10.7710/2162-3309.1035/> (accessed August 18, 2016).
- Muilenberg, J., Lebow, M. and Rich, J. (2014), "Lessons learned from a research data management pilot course", *Journal of eScience Librarianship*, Vol. 3 No. 1, pp. 67-73.
- Myers, J., Allison, T.C., Bittner, S., Didier, B., Frenklach, M., Green, W.H. Jr, Ho, Y., Hewson, J., Koegler, W., Lansing, C., Leahy, D., Lee, M., McCoy, R., Minkoff, M., Nijssure, S., von Laszewski, G., Montoya, D., Oluwole, L., Pancerella, C., Pinzon, R., Pitz, W., Rahn, L.A., Ruscic, B., Schuchardt, S., Stephan, E., Wagner, A., Windus, T. and Yanget, C. (2005), "A collaborative informatics infrastructure for multi-scale science", *Cluster Computing*, Vol. 8 No. 2, pp. 244-253.
- National Academy of Science (2009), *Ensuring the Integrity, Accessibility, and Stewardship of Research Data*, National Academy of Science, Washington, DC.
- National Research Council of the National Academies (2015), *Preparing the Workforce for Digital Curation*, National Academies Press, Washington, DC.
- National Science Board (2005), *Long-Lived Data Collections*, National Science Board, Arlington, VA.
- Nature (2009), "Data's shameful neglect", *Nature*, September 10, p. 145.
- Nicholls, N., Samuel, S.M., Lalwani, L.N., Grochowski, P.F. and Green, J.A. (2014), "Resources to support faculty writing data management plans: lessons learned from an engineering pilot", *International Journal of Digital Curation*, Vol. 9 No. 1, pp. 242-252.
- Nicholson, S.W. and Bennett, T.B. (2011), "Data sharing: academic libraries and the scholarly enterprise", *Portal: Libraries and the Academy*, Vol. 11 No. 1, pp. 505-516.
- Noonan, D. and Chute, T. (2014), "Data curation and the university archives", *American Archivist*, Vol. 77 No. 1, pp. 201-240.
- Organisation for Economic Co-Operation and Development (2007), *OECD Principles and Guidelines for Access to Research Data from Public Funding*, Organisation for Economic Co-Operation and Development, Paris.
- Osswald, A. (2013), *Skills for the Future: Educational Opportunities for Digital Curation Professionals*, European Commission, Florence.
- Ovadia, S. (2013), "Digital content curation and why it matters to librarians", *Behavioral and Social Sciences Librarian*, Vol. 32 No. 1, pp. 58-62.

- Palmer, C.L. and Cragin, M.H. (2008), "Scholarship and disciplinary practices", *Annual Review of Information Science and Technology*, Vol. 42 No. 1, pp. 165-212.
- Palmer, C.L., Teppeau, L. and Pirmann, C.M. (2009), *Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development*, OCLC Research, Dublin, OH.
- Palmer, C.L., Weber, N.M. and Cragin, M.H. (2011), *The Analytic Potential of Scientific Data: Understanding Re-Use Value*, ASSIST, New Orleans.
- Parsons, M. and Duerr, R. (2005), "Designating user communities for scientific data: challenges and opportunities", *Data Science Journal*, Vol. 4, pp. 31-38.
- Parsons, M. and Fox, P. (2013), "Is data publication the right metaphor?", *Data Science Journal*, Vol. 12, pp. WDS34-WDS43.
- Patrick, M. and Wilson, J. (2013), "Getting data creators on board with the digital curation agenda", *Proceedings of the Framing the Digital Curation Curriculum Conference*, European Commission, Florence.
- Polanyi, M. (1966), *The Tacit Dimension*, Anchor Books, Garden City.
- Poole, A.H. (2015), "How has your science data grown? Digital curation and the human factor, a critical literature review", *Archival Science*, Vol. 15 No. 2, pp. 1-39.
- Preservation and Metadata Working Group (n.d.), *DataONE Preservation Strategy*, National Science Foundation, Washington, DC.
- Price, L. and Smith, A. (2000), *Managing Cultural Assets from a Business Perspective*, Council on Library and Information Resources, Washington, DC.
- Prom, C. (2011), "Making digital curation a systematic institutional function", *International Journal of Digital Curation*, Vol. 6 No. 1, pp. 139-152.
- Pryor, G. (2009), "Multi-scale data sharing in the life sciences: some lessons for policy makers", *International Journal of Digital Curation*, Vol. 4 No. 3, pp. 71-82.
- Pryor, G. (2012), "Why manage research data?", in Pryor, G. (Ed.), *Managing Research Data*, Facet, London, pp. 1-16.
- Pryor, G. (2013), "A maturing process of engagement: raising data capabilities in UK higher education", *International Journal of Digital Curation*, Vol. 8 No. 2, pp. 181-193.
- Ray, J. (2009), "Sharks, digital curation, and the education of information professionals", *Museum Management and Curatorship*, Vol. 24 No. 4, pp. 357-368.
- Ray, J. (2014), "Introduction to research data management", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 1-23.
- Redwine, G., Barnard, M., Donovan, K., Farr, E., Forstrom, M., Hansen, W., John, J.L., Kuhl, N. and Shaw, S. (2013), *Born Digital: Guidance for Donors, Dealers, and Archival Repositories*, Council on Library and Information Resources, Washington, DC.
- Reilly, J.B.F. and Waltz, M.E. (2014), "Trustworthy data repositories", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 109-126.
- Repository Task Force (2009), *The Research Library's Role in Digital Repository Services: The Final Report of the ARL Digital Repository Issues Task Force*, Association of Research Libraries, Washington, DC.
- Riley, J. (2014), "Metadata services", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 149-165.



- Rothenberg, J. (1995), "Ensuring the longevity of digital documents", *Scientific American*, Vol. 272 No. 1, pp. 42-47.
- Rusbridge, C. (2007), "Create, curate, re-use: the expanding life course of digital research", Educause Australia, Melbourne, pp. 1-11.
- Rusbridge, C., Burnhill, P., Ross, S., Buneman, P., Giaretta, D. and Lyon, L. (2005), *The Digital Curation Centre: A Vision for Digital Curation*, IEEE Computer Society, Washington, DC, pp. 31-41.
- Sallans, A. and Donnelly, M. (2012), "DMP online and DMPTool: different strategies towards a shared goal", *International Journal of Digital Curation*, Vol. 7 No. 2, pp. 123-129.
- Sallans, A. and Lake, S. (2014), "Data management assessment and planning tools", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 87-107.
- Sandusky, R., Palmer, C.L., Allard, S., Cragin, M.H., Cruse, P., Renear, A. and Tenopir, C. (2009), "The DataNet partners: sharing science, linking domains, curating data", *Proceedings of the American Society for Information Science and Technology*, Vol. 46 No. 1, pp. 1-8.
- Sayogo, D. and Pardo, T.A. (2013), "Exploring the determinants of scientific data sharing: understanding the motivation to publish research data", *Government Information Quarterly*, Vol. 30, pp. S19-S31.
- Saramozzino, J.M., Ramirez, M.L. and McGaughey, K.J. (2012), "A study of faculty data curation behaviors and attitudes at a teaching-centered university", *College & Research Libraries*, Vol. 73 No. 4, pp. 349-365.
- Shaffer, J. (2013), "The role of the library in the research enterprise", *Journal of E-Science Librarianship*, Vol. 2 No. 1, pp. 8-15.
- Shankar, K. (2007), "Order from chaos: the poetics and pragmatics of scientific recordkeeping", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 10, pp. 1457-1466.
- Shorish, Y. (2012), "Data curation is for everyone! the case for master's and baccalaureate institutional engagement with data curation", *Journal of Web Librarianship*, Vol. 6 No. 4, pp. 263-273.
- Smith, M. (2014), "Data governance: where technology and policy collide", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 45-59.
- Starr, J., Willett, P., Federer, L., Horning, C. and Bergstrom, M.L. (2012), "A collaborative framework for data management services: the experience of the University of California", *Journal of eScience Librarianship*, Vol. 1 No. 2, pp. 109-114.
- Steinhart, G. (2014), "An institutional perspective on data curation services", in Ray, J. (Ed.), *Research Data Management: Practical Guidance for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 303-323.
- Steinhart, G., Chen, E., Arguillas, F., Dietrich, D. and Kramer, S. (2012), "Prepared to plan? A snapshot of researcher readiness to address data management planning requirements", *Journal of eScience Librarianship*, Vol. 1 No. 2, pp. 63-78.
- Strasser, C. (2015), *Research Data Management*, National Information Standards Organization, Baltimore.
- Strasser, C., Abrams, S. and Cruse, P. (2014), "DMPTool 2: expanding functionality for better data management planning", *International Journal of Digital Curation*, Vol. 9 No. 1, pp. 324-330.
- 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*, Key Perspectives, Ltd, Truro.

- Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A.U., Wu, L., Read, E., Manoff, M. and Frame, M. (2011), "Data sharing by scientists: practices and perceptions", *PLoS ONE*, Vol. 6 No. 6, pp. 1-20.
- Tenopir, C., Birch, B. and Allard, S. (2012), *Academic Libraries and Research Data Services: Current Practices and Plans for the Future*, ACRL, Chicago, IL.
- Tenopir, C., Levine, K., Allard, S., Christian, L., Volentine, R., Boehm, R., Nichols, F., Nicholas, D., Jamali, H.R., Herman, E. and Watkinson, A. (2015), "Trustworthiness and authority of scholarly information in a digital age: results of an international questionnaire", *Journal of the Association for Information Science and Technology*, pp. 1-14.
- Tibbo, H.R. (2015), "Digital curation education and training: from digitization to graduate curricula to MOOCs", *International Journal of Digital Curation*, Vol. 10 No. 1, pp. 144-153.
- Treloar, A., Choudhury, G.S. and Michener, W. (2012), "Contrasting national research data strategies: Australia and the USA", in Pryor, G. (Ed.), *Managing Research Data*, Facet, London, pp. 173-203.
- Van den Eynden, V., Corti, L., Woollard, M., Bishop, L. and Horton, L. (2010), *Managing and Sharing Data: Best Practices for Researchers*, UK Data Archive, Essex.
- Vivarelli, M., Cassella, M. and Valacchi, F. (2013), *The Digital Curator Between Continuity and Change: Developing a Training Course at the University of Turin*, European Union, Florence.
- Wallis, J., Borgman, C., Mayernik, M. and Pepe, A. (2008), "Moving archival practices upstream: an exploration of the life cycle of ecological sensing data in collaborative field research", *International Journal of Digital Curation*, Vol. 1 No. 3, pp. 114-126.
- Wallis, J.C., Rolando, E. and Borgman, C.L. (2013), "If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology", *PLoS One*, Vol. 8 No. 7, pp. 1-17.
- Walters, T.O. (2009), "Data curation program development in US universities: the Georgia Institute of Technology example", *International Journal of Digital Curation*, Vol. 4 No. 3, pp. 83-92.
- Walters, T.O. (2014), "Assimilating digital repositories into the active research process", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 189-201.
- Walters, T.O. and Skinner, K. (2011), *New Roles for New Times: Digital Curation for Preservation*, Association of Research Libraries, Washington, DC.
- Wang, Z. (2013), "Co-curation: new strategies, roles, services, and opportunities for libraries in the post-web era and the digital media context", *Libri*, Vol. 63 No. 2, pp. 71-86.
- Waters, D.J. (2007), "Doing much more than we have so far attempted", *EduCAUSE Review*, Vol. 42 No. 5, pp. 8-9.
- Wenger, E. (2006), "Communities of practice: a brief introduction", available at: <https://scholarsbank.uoregon.edu/xmlui/bitstream/handle/1794/11736/A%20brief%20introduction%20to%20CoP.pdf?sequence=1> (accessed August 15, 2016).
- Westra, B. (2014), "Developing data management services for researchers at the university of Oregon", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 375-391.
- Whitlock, M.C. (2011), "Data archiving in ecology and evolution: best practices", *Trends in Ecology and Evolution*, Vol. 26 No. 2, pp. 61-65.
- Whyte, A., Molloy, L., Beagrie, N. and Houghton, J. (2014), "What to measure? Toward metrics for research data management", in Ray, J. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 275-300.

- Winget, M., Frick, C., McDonough, J., Renear, A. and Lowood, H. (2009), *Digital Curation of Humanistic, Multimedia Materials: Lessons Learned and Future Directions*, School of Information and Library Science, University of North Carolina, Chapel Hill, NC, pp. 42-43.
- Witt, M. (2008), "Institutional repositories and research data curation in a distributed environment", *Library Trends*, Vol. 57 No. 2, pp. 191-201.
- Witt, M., Carlson, J., Brandt, D.S. and Cragin, M. (2009), "Constructing data curation profiles", *International Journal of Digital Curation*, Vol. 4 No. 3, pp. 93-103.
- Wood, D.J. and Gray, B. (1991), "Toward a comprehensive theory of collaboration", *Journal of Applied Behavioral Science*, Vol. 27 No. 2, pp. 139-162.
- Wright, S., Whitmire, A., Zilinski, L. and Minor, D. (2014), "Collaboration and tension between institutions and units providing data management support", *Bulletin of the American Society for Information Science and Technology*, Vol. 40 No. 6, pp. 18-21.
- Yakel, E. (2007), "Digital curation", *OCLS Systems and Services*, Vol. 23 No. 4, pp. 338-339.
- Yarmey, L. and Baker, K.S. (2013), "Towards standardization: a participatory framework for scientific standard-making", *International Journal of Digital Curation*, Vol. 8 No. 1, pp. 157-172.
- Zhang, T., Zilinski, L., Brandt, D.S. and Carlson, J. (2015), "Assessing perceived usability of the data curation profile toolkit using the technology acceptance model", *International Journal of Digital Curation*, Vol. 10 No. 1, pp. 48-67.
- Zimmerman, A. (2008), "New knowledge from old data: the role of standards in the sharing and reuse of ecological data", *Science, Technology & Human Values*, Vol. 33 No. 5, pp. 631-652.
- Zorich, D. (2009), *Working Together or Apart: Promoting the Next Generation of Digital Scholarship*, Council on Library and Information Resources, Washington, DC.

#### Further reading

- National Science Foundation (2005), *NSF's Cyberinfrastructure Vision for 21st Century Discovery*, National Science Foundation, Arlington, VA.

#### Corresponding author

Alex H. Poole can be contacted at: [ahp56@drexel.edu](mailto:ahp56@drexel.edu)

---

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldgroupublishing.com/licensing/reprints.htm](http://www.emeraldgroupublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)