"Natural allies"

Librarians, archivists, and big data in international digital humanities project work

Alex H. Poole

*Department of Information Science, Drexel University,
Philadelphia, Pennsylvania, USA, and*

Deborah A. Garwood

*College of Computing and Informatics, Drexel University, Philadelphia,
Pennsylvania, USA*

Abstract

Purpose – In Digging into Data 3 (DID3) (2014-2016), ten funders from four countries (the USA, Canada, the UK, and the Netherlands) granted $5.1 million to 14 project teams to pursue data-intensive, interdisciplinary, and international digital humanities (DH) research. The purpose of this paper is to employ the DID3 projects as a case study to explore the following research question: what roles do librarians and archivists take on in data-intensive, interdisciplinary, and international DH projects?

Design/methodology/approach – Participation was secured from 53 persons representing eleven projects. The study was conducted in the naturalistic paradigm. It is a qualitative case study involving snowball sampling, semi-structured interviews, and grounded analysis.

Findings – Librarians or archivists were involved officially in 3 of the 11 projects (27.3 percent). Perhaps more importantly, information professionals played vital unofficial roles in these projects, namely as consultants and liaisons and also as technical support. Information and library science (ILS) expertise helped DID3 researchers with issues such as visualization, rights management, and user testing. DID3 participants also suggested ways in which librarians and archivists might further support DH projects, concentrating on three key areas: curation, outreach, and ILS education. Finally, six directions for future research are suggested.

Originality/value – Much untapped potential exists for librarians and archivists to collaborate with DH scholars; a gap exists between researcher awareness and information professionals' capacity.

Keywords Collaboration, Interdisciplinarity, Digital humanities, Data management, Digital curation, Data curation, Information and library science

Paper type Research paper

Introduction

There is no shortage of enthusiasm and optimism in discussions of digital humanities (DH). Klein (2015) observes, "Digital Humanities is a rapidly growing field at the intersections of computing and the disciplines of humanities and arts, interdisciplinary fields of culture and communication, and the professions of education and library and information science" (pp. 1-2). DH appears "a vibrant and rapidly growing field of endeavor" to Schreibman *et al.* (2016, p. xvii) and a "quickly evolving, contested and exciting field" to Svensson and Goldberg (2015, p. 1). Pannapacker (2012) boldly claims, "At this point, the digital humanities are 'the thing.' There's no "next" about it. And it will not be long until the digital humanities are, quite simply, 'the humanities'" (p. 233). New publications, organizations, networks, research centers, academic programs, and funded projects attest to its robustness (Klein, 2015).

Exploiting these channels, conversations in DH scholarship concentrate on a bevy of topics: historicizing, defining, conceptualizing, theorizing, analyzing, critiquing, criticizing, creating, building, producing, practicing, collaborating, disciplining and interdisciplining, developing infrastructure, teaching, institutionalizing, professionalizing, disseminating, rewarding, and archiving (Bartsherer and Coover, 2011; Berry, 2012; Burdick *et al.*, 2012; Cohen and Scheinfeldt, 2013; Crompton *et al.*, 2016; Deegan and McCarty, 2012; Gardiner and Musto, 2015; Gold, 2012; Jones, 2014; Klein, 2015; Klein and Gold, 2016; Nyhan *et al.*, 2013;

Schreibman *et al.*, 2004; Schreibman *et al.*, 2016; Siemens and Schreibman, 2007; Svensson
and Goldberg, 2015; Warwick *et al.*, 2012).

The Digging into Data Challenges, first offered in 2009, constitute one of the most
important endeavors in this flurry of activity. Its first call for proposals announced,
"The idea behind the Digging into Data Challenge is to answer the question 'what do you
do with a million books?' Or a million pages of newspaper? Or a million photographs of
artwork? That is, how does the notion of scale affect humanities and social science
research?"[1] Digging into Data 1 stimulated innovative analytical research techniques
involving large data sets, and encouraged interdisciplinary, international collaboration,
and partnering with data repositories to ensure efficient access to their materials
(See footnote 1). Eight international teams received awards from four funders based in
three countries (Table AI).

These projects broke new ground. According to a Williford and Henry (2012) report,
the first Digging into Data Challenge introduced "a new paradigm: a digital ecology of
data, algorithms, metadata, analytical and visualization tools, and new forms of scholarly
expression that result from this research" (Williford and Henry, 2012, p. 2). Digging into
Data 2 featured eight funders from four nations; 13 projects received funding.

At the end of 2013, ten funders from four countries (the USA, Canada, the UK, and
the Netherlands) awarded a total of $5.1 million to fourteen international collaborations
comprising scholars, scientists, and information professionals. This article employs
these projects from Digging into Data 3 (DID3) as a case study to explore the following
research question:

RQ1. What roles and responsibilities do information professionals such as librarians and
archivists assume in data-intensive, interdisciplinary, and international DH projects?

This paper first reviews the literature, concentrating on the relationships among DH,
Library and Information and Science (LIS), and libraries and librarians. Second, it explains
and justifies the study's qualitative approach. Third, it reports the findings of the study and
discusses their ramifications, focusing on librarians' and archivists' official and unofficial
involvement in projects and on their specific roles and responsibilities. Additionally,
it probes three issues that speak directly to bolstering librarians' and archivists'
participation in projects such as DID3: digital curation, LIS education and professional
preparation, and outreach opportunities for librarians and archivists. Fifth, conclusions and
six recommendations for future research are advanced.

Literature review
Scholars discuss the ways in which DH and LIS complement each other, the possible roles
libraries and librarians can play in DH work, and the challenges libraries and librarians face
when engaging in such work.

DH and library and information science
As Bawden and Robinson (2012), Dalbello (2011), Koltay (2016), Poole (2017a, b), Robinson
et al. (2015), and White and Gilbert (2016) suggest, DH and information and library science
(ILS) complement each other. First, both fields focus on recorded information and
documents. Second, both concentrate in research and in practice on collection development
and management; searching and retrieval; digital libraries, archives, and repositories;
metadata and description; ontologies, taxonomies, and classification; scholarly production
and dissemination; open access and linked data; bibliography and bibliometrics;
digitization; preservation and curation; user experience; interfaces and browsing;
visualization; and Big Data and data mining (Robinson *et al.*, 2015). Third, both
developed from service functions and often occupy the same academic units – even the same

physical spaces. Fourth, both include topics from the other in their curricula and publish in each other's venues. Fifth, both embrace interdisciplinarity. Sixth, both have been influenced by the data-intensive fourth paradigm, which embraces technologies and techniques enabling data-intensive scientific exploration (Bell *et al.*, 2009). Finally, both fields appear anxious about their future in the academy. In short, both fields rest on a theoretical and conceptual basis that underpins their practical activity (Robinson *et al.*, 2015). This literature suggests clear potential for symbiosis.

Just as scholars see DH and LIS as complementary, so they argue that DH and libraries constitute natural partners (Bryson *et al.*, 2011; Cunningham, 2010; Elliott *et al.*, 2015; Rockenbach, 2013; Sula, 2013). Keener (2015) frames libraries and DH as "inextricably linked" and Vandegrift and Varner (2013) sees them "evolving in common." Indeed, an Association of Research Libraries study of 2011 found libraries hard-pressed to keep up with escalating demand for DH services (Bryson *et al.*, 2011).

These assertions do not surprise. Fundamentally, libraries complement DH because they pivot around collecting, cataloging, describing, organizing, providing access to, facilitating the use of, storing, protecting, preserving, and repurposing data (Courtney and Dallis, 2015; Elliott *et al.*, 2015; Fay and Nyhan, 2015; Kamada, 2010; Moazeni, 2015; Schaffner and Erway, 2014; Vandegrift and Varner, 2013; White, 2016). In this sense, all research libraries currently support DH work (Schaffner and Erway, 2014). Bryson *et al.*'s (2011) sample of librarians evinced a vested interest for DH projects to be closely tied to their libraries (Bryson *et al.*, 2011).

Roles for librarians and archivists
Scholarly conversations around role and responsibilities for librarians and archivists revolve around their being partners or supporters (or both) of DH work and the range of roles and responsibilities they may assume in such work.

Partners vs supporters
Many scholars continue to view the library as a purchaser and a repository (Jahnke and Asher, 2012; Munoz and Renear, 2011). They relegate librarians to service providers. Research, by contrast, remains the intellectual bailiwick of faculty (Munoz, 2016). Notably, DH scholars rarely look to librarians for domain expertise and only occasionally do librarians comprise full-fledged participants of DH project teams (Schaffner and Erway, 2014).

Rejecting this division – not to say hierarchy – of labor, Munoz (2016) argues in favor of DH in the library as an intellectual wellspring and Sacco and her colleagues (2015) urge librarians to create new scholarly knowledge. Vandegrift and Varner (2013) meanwhile encourage librarians to be "true scholarly partner[s]" with digital humanists (p. 70), as does Keener (2015) and Brandenburg *et al.* (2017). In the same spirit, Posner (2013) advocates for librarians to assume the role of collaborators as opposed to supporters; she insists, "DH expertise is […] not a service to be offered in silent support of a scholar's master plan" (p. 46). Tenopir *et al.* (2015) even see librarians as potential partners "in all aspects of the research process, from data collection to publication and preservation of research output" (Tenopir *et al.*, 2015, p. 2). It remains unclear, however, how much original DH research librarians initiate – much less what value scholars attribute to such efforts (Cunningham, 2010).

The purported dichotomy between research and support appears overdrawn, however. For instance, Varner (2015) favors invigorating the library as a locus of production and storage. Rockenbach (2013) takes a parallel tack, lobbying for librarians to serve as partners and collaborators even as she urges librarians not to eschew their service roles altogether. Sacco *et al.* (2015) and Zhang *et al.*, (2015) too, see the role of partner and that of supporter as symbiotic.

Lamenting the "arguably unproductive dichotomy of librarians as collaborators vs supporters," Keener (2015) concludes, "Librarians and staff can provide ad hoc DH services

while also being research partners." Initial consultations may spur deeper and lengthier collaborations (Keener, 2015), even though such collaborations remain challenging to inaugurate as well as to sustain (Zhang *et al.*, 2015).

Range of involvement
Librarians' roles in DH projects appear largely ad hoc (Bryson *et al.*, 2011). Posner (2013) characterizes library involvement in DH as consisting of "a 'center,' a 'suite of services,' a librarian with a revised job title, or, murkiest of all, an 'initiative'" (p. 43) and Sula (2013) and Schaffner and Erway (2014) find that libraries' DH initiatives remain motley, even eclectic. Most libraries mold their DH initiatives to fit local circumstances (Elliott *et al.*, 2015). Other scholars exhort libraries to pursue a robust, scalable, and sustainable DH model (Keener, 2015; Vinopal and McCormick, 2013).

Despite these differing interpretations, scholars identify eight primary roles for librarians and archivists in DH (Table I). The role of digital curator, which involves adding value to data to promote sharing and reuse over its lifecycle and includes digital preservation and management, takes on particular salience. Digital curation offers opportunities for academic libraries to embed themselves further in the research lifecycle (Cox *et al.*, 2017; Tenopir *et al.*, 2012).

Challenges for libraries and librarians
Challenges facing librarians in DH stem from policies, resources and infrastructure, outreach and awareness, and organizational culture.

Policies. First, a lack of settled policies, protocols, and procedures trammel librarians and DH scholars alike (Bryson *et al.*, 2011). Collecting, description, access, and preservation policies in particular exert a powerful influence on DH work (Sacco *et al.*, 2015; White, 2016). Policies also impact resource allocation and infrastructure development and maintenance. A lack of policy lacunae may jeopardize the sustainability of librarians' DH involvement.

Resources and infrastructure. Second, DH work in the library confronts the perennial challenges of resources, primarily staff time, staff training resources, and funding (Bryson *et al.*, 2011; Courtney and Dallis, 2015; Posner, 2013). Libraries' administrative and

Roles for librarians	References in relevant literature
Support pedagogy	Pun (2015), Varner (2015), White (2016), Zhang *et al.* (2015)
Provide physical/virtual space	Courtney and Dallis (2015), Cunningham (2010), Elliott *et al.* (2015), Rockenbach (2013), Schaffner and Erway (2014), White and Gilbert (2016), Williford and Henry (2012), Zhang *et al.* (2015)
Resource discovery and provision of access/dissemination (electronic/born-digital/digitized or physical materials)	Cunningham (2010), Elliott *et al.* (2015), Kamada (2010), Keener (2015), Rockenbach (2013), Schaffner and Erway (2014), Showers (2012), Sula (2013), Vandegrift and Varner (2013), Williford and Henry (2012)
Metadata	Kamada (2010), Keener (2015), Schaffner and Erway (2014), Showers (2012), Sula (2013), Tenopir *et al.* (2012)
Project management	Elliott *et al.* (2015), Munoz (2016), Williford and Henry (2012)
Consultations/instruction	Elliott *et al.* (2015), Keener (2015), Showers (2012), Sula (2013), Tenopir *et al.* (2012), Williford and Henry (2012), Zhang *et al.* (2015)
Sustainability/hosting/storage	Kretzschmar and Gray (2010), Schaffner and Erway (2014), Showers (2012), Tenopir *et al.* (2012), Williford and Henry (2012)
Digital preservation/curation	Cox *et al.* (2017), Flanders and Munoz (2016), Gold (2010), Henry (2014), Kamada (2010), Poole (2013), Poole (2017b), Sabharwal (2015), Poole (2017a), Sabharwal (2017), Schaffner and Erway (2014), Showers (2012), Sula (2013), Tenopir *et al.* (2012, 2014, 2015), Williford and Henry (2012), Zhang *et al.* (2015)

Table I.
Roles for librarians in
digital humanities

technical infrastructures undergird DH, moreover, but both remain nascent for DH work (Posner, 2013; White, 2016). Finally, the very eclecticism of DH projects renders it difficult for libraries to scale their support (Elliott *et al.*, 2015; Vinopal and McCormick, 2013).

Outreach and raising awareness. Third, whether by intention or ignorance, many DH scholars neglect to consult librarians. Indeed, their existing attitudes toward library services influence their tendency to seek out or avoid the library for DH support (Schaffner and Erway, 2014). To promote awareness and effect outreach, librarians may facilitate cross-disciplinary engagement among researchers and students at libraries through workshops, or outside of libraries by hosting events (Keener, 2015; Williford and Henry, 2012; Zhang *et al.*, 2015).

Organizational culture. Fourth, organizational culture "focuses on beliefs, values, and meanings used by members of an organization, and the practices and behaviors that exemplify and reinforce them" (Lakos and Phipps, 2004, p. 348). Culture immanently resists change, but libraries need an "inquisitive, adaptable, [and] responsive" organizational culture to optimize their role in DH (Vinopal and McCormick, 2013, p. 40). Rockenbach (2013) states, "DH is messy. It involves uncertainty, deep collaborations, and a flexibility that is foreign to traditional library culture" (p. 6). Perhaps even more daunting, DH in the library may muddle existing workflows and undercut traditional performance metrics (Posner, 2013).

What is more, DH demands stalwart leadership from library administrators (Lakos and Phipps, 2004; Rockenbach, 2013). Responsibilities for leaders include "articulating a strategic vision, defining priorities, addressing the connections between new services and established ones, facilitating horizontal as well as vertical communication and collaboration, and building a staff that are lifelong learners with evolving job descriptions" (Vinopal and McCormick, 2013, pp. 40-41).

Momentum for DH investment must come from the grassroots, too. Librarians facing DH opportunities may themselves struggle with undue cautiousness-cum-modesty (Posner, 2013; Sacco *et al.*, 2015) or even timidity born of an "academic inferiority complex" (Vandegrift and Varner, 2013, p. 76). Tacit or explicit hierarchies may also sap librarians' efforts to collaborate with faculty or graduate students (Keener, 2015; Posner, 2013). The best strategy, therefore, remains to frame DH in the context of the library's overall mission and policies (Elliott *et al.*, 2015; Kretzschmar and Gray, 2010).

Methods

This study focuses on the most recent round (2014-2016) of the Digging into Data projects to maximize the likelihood of interviewee recall. Participation by representatives from 11 of the 14 projects (53 persons) from a variety of disciplines and organizational units and with a variety of job titles and associated roles and responsibilities was secured (Tables II-IV). (Drexel University IRB 1601004139). Interviews were conducted between September and December of 2016 by phone, Skype, and Google Hangouts; each one lasted between 30 and 90 minutes and all were recorded and transcribed.

This study embraces the naturalistic paradigm, which concentrates on describing the characteristics of and understanding a single social phenomenon (Mellon, 1990). It is a qualitative case study based on a human instrument, semi-structured interviews, purposive sampling, and grounded data analysis (Lincoln and Guba, 1985). Focusing on exploration, description, and comparison (Bernard and Ryan, 2010; Gorman and Clayton, 2005; Sutton, 1993; Taylor and Bogdan, 1998; Westbrook, 1994), qualitative methods offer "richer opportunities for gathering and assessing [...] what the participant values, believes, thinks, and feels about social life" (Saldana, 2013, p. 92).

Qualitative case studies enable the description and understanding of nuanced, complex social phenomena. They permit an in-depth and real-life focus on contemporary events

Professor	13
Manager	9
Postdoctoral researcher	7
PhD student/candidate	6
Associate professor	4
Assistant professor	4
Director/associate director/assistant director	3
Lecturer	1
Senior lecturer	1
Senior researcher	1
Research chair	1
Biodiversity specialist	1
Computer media specialist	1
Data analyst	1
Total	53

Table II.
Participants: job titles

Computer science	8
Languages, literatures, linguistics	7
Management/business	6
Biodiversity/bioinformatics	4
Communication studies/arts	4
History/historical research	4
East asian studies	3
English	3
Political science	2
Visualization	2
Anthropology	1
Cultural analytics	1
Deafness, cognition and language	1
Design studies department	1
Digital humanities	1
Industrial and organizational psychology	1
Mathematics and natural sciences	1
Medieval and early modern studies	1
Spatial and textual analysis	1
Text mining	1
Total	53

Table III.
Participants:
disciplinary affiliation

Project number	PIs per project interviewed	Other project personnel interviewed	Total number of participants interviewed
01	2	5	7
02	1	2	3
03	2	1	3
04	2	1	3
05	1	2	3
06	1	0	1
07	2	6	8
08	2	5	7
09	3	4	7
10	2	6	8
11	1	2	3
Total	19	34	53

Table IV.
Participants
by project

when behaviors resist manipulation and measurement and when the boundaries between phenomenon and context remain porous. Nonetheless, case studies are only theoretically generalizable (Bernard and Ryan, 2010; Choemprayong and Wildemuth, 2009; King, 1994; Sholtz and Tietje, 2002; Yin, 2009).

Semi-structured interviews allow an investigator to integrate multiple perspectives, to discern processes, to develop holistic descriptions, and to reconstruct events (Rubin and Rubin, 2005; Weiss, 1994). They combine flexibility and control: the interviewer and the participants co-construct the conversation (Charmaz, 2014). Supporting such interviews, snowball sampling allows the recruitment of participants where few in number or difficult to locate (Atkinson and Flint, 2004; Morgan, 2008). First, we contacted DID3 PIs. We then interviewed them, the team members to whom they referred us, or both. Sampling continued until the data became repetitive (saturation) (Pickard, 2013).

Documentary evidence complements interview data. Pivotal in corroborating or verifying evidence from other sources, documents exist in non-reactive form. They vouchsafe historical insight, offer information possibly unavailable in other forms, and potentially give a more accurate representation than data gathered through self-reporting (Hodder, 2000; Wildemuth, 2009; Yin, 2009). Documents consulted in this study included DID3 applications, progress reports, White Papers, and project outputs such as journal articles, press releases, and websites.

A grounded, inductive, and idiographic approach steers data analysis. As Bernard and Ryan (2010) assert, "The less we know about a research problem, the more important it is to take an inductive approach" (p. 266). More specifically, an inductive approach pays dividends because:

> [It] is more likely to identify the multiple realities to be found in those data; because such analysis is more likely to make the investigator-respondent (or object) interaction explicit, recognizable, and accountable; because this process is more likely to describe fully the setting and to make decisions about transferability to other settings easier; because inductive data analysis is more likely to identify the mutually shaping influences that interact; and because values can be an explicit part of the analytic structure (Lincoln and Guba, 1985, p. 40).

Interviews were open coded (initial coding followed by focused coding) by hand and with NVivo (Bazeley, 2013; Saldana, 2013). Data analysis rests upon Jick's (1979) conception of triangulation. Units of analysis included individuals, project teams, projects, and project outputs.

Results and discussion
Findings revolve around the general involvement of librarians and archivists. Participants discuss how and why librarians and archivists were and were not involved both officially and unofficially.

Involvement of librarians and archivists
Even though DID3 enjoined the involvement of librarians and archivists, these projects saw surprisingly few of them officially involved. Persons trained as librarians or archivists took part officially in 3 of the 11 projects (27.3 percent). Four participants (P15-10, P28-08, P41-10, and P49-09) had earned advanced degrees or certifications as librarians or archivists (three MSLIS holders and one certified archivist), two of whom (P15-10 and P41-10) worked in libraries during the project. A fifth participant, P2-09, trained in the humanities worked in a library during DID3. Overall, five of 53 participants (9.4%) were credentialed librarians or worked in libraries during the projects or both.

As striking as the low numbers of librarians and archivists involved, the librarian who played the most foundational role in any project was not a credentialed LIS professional.

Though she worked as a librarian, she had earned a PhD in the humanities. She recalled her "secret agenda" in DID3: to encourage the project team to adopt recent library technology facilitating image interoperability. In so doing, she leveraged her library's robust infrastructure on behalf of the project. P2-09's project used – and in the process enriched – repository data that was then amenable to reuse as an institutional asset. The project thereby showed considerable potential for librarians' involvement with mutually beneficial outcomes for the DID3 project and the university's library alike.

As noted above, four other persons credentialed as librarians or archivists were members of project teams but none of their job titles included "librarian" or "archivist." Nonetheless, their LIS backgrounds proved valuable. For instance, P28-08 parlayed her LIS skills into her work as project manager. P51-08 noted of P28-08's progressive engagement:

> Initially at [home institution], we hired someone to do applied research financial help and project management help and the person we hired happened to have a master's in information sciences […] she started participating in our weekly meetings to help us track from a project management standpoint what was next, what were all the balls we had in the air, and then would contribute more along the side of, "hey, in terms of taxonomy you might also want to consider NISO standards […]" So that became much more of a focus for us. She would bring in, "well, there's someone who's key, Canada's king of data […] You need to talk to him.

Complementing P51-08, P28-08 reflected on the evolution of her own project involvement:

> I've kind of taken on the role of writing the data management part for our research grants. Thinking about, OK, you have all this data; how are you going to preserve it? How are you going to share it? What are you going to do five years from now when the project ends? What are your plans for the data? That's where I've started writing a lot, thinking about things like, OK, what is copyrighted, what is sharable, what is open access. That's really been helpful, just thinking about it from all the different ways we could preserve data. That's something I brought from LIS on that data management side.

In much the same spirit, P28-08 found skills developed in her LIS work useful in evaluating information. She asked rhetorically, "How do you make sure that it is authentic, real, actual credible data or use articles in blog posts […] especially with all this fake news going around? That's where I feel like that skill has really come in."

Contributions of librarians and archivists

The apparent non-involvement of librarians and archivists in these DID3 projects was quite deceptive, perhaps because it was elemental. Of her PI, for example, P46-10 observed, "He was the person bringing the library-like expertise, but I'm not sure that would qualify as a librarian. He was a librarian to me in the project." In other words, project team members who had neither the MLIS credential nor a job title that included the term librarian or archivist nonetheless employed skills and competencies rooted in LIS. Similarly, P44-07, a History PhD, claimed librarians were not involved in her project, but then noted that she worked at an institute that in effect is a library.

In sum, librarians played key roles in DID3 projects, albeit unofficially, and largely as supporters rather than full partners. Concerning librarians at her institution, PI P1-01 commented, "I had some informal discussions, but nothing really elaborate." Likewise, PI P47-08's team consulted librarians "here and there." Finally, P28-08 related, "We actually met with some people in LIS, a couple of my former [professors] […] they're not actively involved in the project, but they've provided that feedback, different ways to think about building that taxonomy in the future."

Participants mentioned eight contributions (in no rank order) made by librarians or archivists. First, project teams capitalized on their libraries' physical space for face-to-face meetings. P3-09, for example, recalled her team meeting frequently at the university library.

Second, librarians intermediated between researchers and publishers. P32-03 noted, "[Librarians] were crucial in putting us into contact with the ECCO [Eighteenth-Century Collections Online] people."

Third, librarians counseled project teams on fair use and copyright, namely in terms of data mining. They helped P12-08 retrieve article DOIs initially, though she later leveraged CrossRef for this task. P30-10 also harnessed librarians' expertise regarding intellectual property and copyright: "I did a lot of running around trying to figure out what data we could have access to from the journals," she recalled, "and I got quite a bit of help from the librarians on resolving that." Similarly, librarians and archivists in PI P11-02's project secured data mining rights to particular texts that served as its intellectual backbone.

Fourth, librarians' expertise proved valuable for P25-01's team. "They could do things with data visualizations that we couldn't," she recalled, "and use different methods to understand data."

Fifth, librarians engaged in user testing, as P47-08 noted. PI P32-03 remarked, "Where we interact a lot with librarians is if we do an iteration of our software, and we say, "What do you think of how this is going to work? Is it possible? Would it work?" or whatever, they give us good feedback."

Sixth, librarians such a P2-09 engaged in translation work. She said, "Talking to a scholar about the fact that they've got 100,000 images available by an API seems to mean nothing to certain faculty members, and what you have to do is get them to tell you what sorts of things they might do if they were looking across the corpus, and then start bringing them along in terms of the technical ways that that could be done. So there are some soft skills there that are important."

Seventh, libraries helped make project data visible and available by hosting it. As P23-07 noted, "The archives people […] have helped make the stuff available and point to it on their website and publicize it." She elaborated, "They weren't involved as officially part of the project. It's just that they had the data that we processed and we make it available to them and to the world now."

Eighth, as P16-06 observed, the libraries undertook "long-term preservation and maintenance and all that," i.e. curation responsibilities. In fact, the libraries at her institution first assumed this responsibility in 2011, when the precursor project to DID3 – the springboard for DID3 – developed the corpus that DID3 subsequently reused.

These eight areas of involvement indicate that librarians and archivists were crucial to projects' success. Even if they were not official members of the DID3 project teams, they assumed a crucial support role in almost all cases. As P10-07 observed, archivists were "Much more service providers than they were [involved] in any other way."

Non-involvement of librarians and archivists

Eight of the 11 projects (72.7 percent) examined did not include librarians or archivists as official project members, despite DID3 funders' professed commitment to including information professionals. Institutional readiness (namely resources and infrastructure and policy) and skill sets militated against the involvement of librarians and archivists.

Resources and infrastructure. Addressing institutional readiness, participants viewed resources and infrastructure as barriers to librarians' and archivists' involvement. P45-01 commented, "It can be tricky […] for [librarians] to justify devoting their own staff time to smaller projects. At some point, you have to say they are a limited resource; what can they do [to get the] biggest bang for their buck, basically. Where is that staff time investment going to have the biggest payoff?" She advised, "Sometimes you just have to rely on them being able to do a little knowledge transfer with you to the point where you can take it on your own."

P36-09 called attention to ways in which a lack of resources could negatively impact infrastructure. She observed, "We have a library, but we are a state institution, which means

we don't have an enormous amount of funding. Our digital infrastructure though the library is […] just not robust. The thought of them being able to handle whatever large data set that we are dealing with, I wouldn't have a lot of confidence." Similarly, the configuration of P53-07's institution dictated her data management strategy: neither the institution's research nor undergraduate libraries engaged appreciably in digital research or curation. A member of the same team, but working at another university, P14-07 expressed a parallel opinion: "We wanted to avoid using [the library's] infrastructure; we wanted more flexibility in terms of the design of the website. We wanted to be able to update it, so we didn't want to bring in institutional resources if we didn't have to."

These concerns harken to Posner's (2013) and Cox *et al.*'s (2017) arguments: many universities lack optimal, or even acceptable, technical infrastructure. Therefore, the lack of library readiness may by extension prevent librarians from project participation, even if they possess the requisite skills and knowledge.

Policy. Extending Posner (2013) and Cox *et al.* (2017), another institutional challenge concerning library involvement took root in policy and by implication, in institutional investment in networked data research. PI P1-01 related, "if you really want to work on developing a project, you need a pretty firm institutional policy that is either going to help you or not. But the in-between policies are just a source of frustration." This was indeed the case at her own institution. P5-10 found the lack of institutional policies at her university maddening; she favored policies to address supporting and sustaining grant-created resources in institutional repositories. "It's not up to me; it's up to universities," she lamented.

Data management policies facilitate coherence, consistency, transparency, and accountability; universities increasingly realize their importance, but quite a gap yawns between recognition and development – much less development and implementation (Bryson *et al.*, 2011; Council on Library and Information Resources, 2013; Cox *et al.*, 2017; Levine, 2012; Oliver and Harvey, 2016; White, 2016). Both of these barriers to data management – infrastructure shortcomings and policy lacunae – indicate a need for top-down leadership, an argument advanced by Rockenbach (2013) and Vinopal and McCormick (2013). Librarians alone can scarcely surmount these structural barriers.

Skills. Most notably, DID3 participants thought librarians and archivists lacked requisite skills. "I don't mean this in the wrong way or in a mean way, but my guess is that most of the skills one would need would not be held by most library staff," claimed P12-08. She provided an example: "I really don't think if I went into the library I would have learned from someone who's a staff member […] how to use MongoDB." As P12-08 insinuates, DH skills and tools may be exceptionally specialized whether in terms of the specific project or the specific (inter)discipline or perhaps both. Thus a conundrum presents itself: which ones do librarians need to know?

Other participants cited computational or domain issues or both. Postdoctoral researcher P10-07 observed, "I don't know that we were set up in such a way from the research side, especially the quite heavy computational aspects of what we were doing, and the statistical aspects of it, to involve archivists."

Also critiquing librarians' skills deficit, PI P6-04 reflected, "The librarians that we had at [institution] were not directly useful to us because they hadn't managed linguistic database systems like the kinds that we wanted." Still another participant (P29-05) suggested librarians lacked requisite skills:

I already had a reasonable grounding in what I was looking at, so I knew the sort of places I would find the information I was looking for. So it wasn't a point where I'd have to go to someone and say, "I've been asked to research X. I have no idea about X." It was usually a case of, if I had a question, it would be more specific, and better aimed at someone who was perhaps an expert in that particular field.

A hiring problem plagues university libraries: they "don't hire the type of person who knows what type of research people are doing in the humanities and who has the technical knowledge" (P6-04).

Notwithstanding barriers erected by infrastructure, policy, skills, and resources, librarians and archivists played understated but nonetheless important roles during DID3 projects. By and large they served in an informal advisory capacity. As such, researchers have yet to appreciate the full range of skills and expertise librarians and archivists can bring to bear in data-intensive DH projects such as DID3. P28-08 broached this missed opportunity, stating, "Before [the PIs] started the [DID3 project], if they had done a little of the work on building taxonomy, even talked to someone in LIS, I think they would have started off at a different point that they are now backtracking to." A gap exists between potentiality and actuality, between researchers' understanding of LIS and what LIS could have done for their projects.

Increasing involvement of librarians and archivists
Data from this study suggests ways in which librarians and archivists may become more fully involved in and better prepared for work in collaborative, international, interdisciplinary data-intensive DH projects. Whether or not their projects involved librarians officially, this study's participants engaged three issues with respect to librarianship and archivy: digital curation, Library and Information Science education, and outreach and raising awareness. Each area not only relates to participants' project experience, but also embraces DH scholarship at their institutions more broadly.

Digital curation as a collaborative opportunity
Digital curation in DH comprises the active and ongoing management of data – whether digitized or born-digital – over its entire lifecycle to make it as highly-functional as possible, especially for sharing and reuse (Flanders and Munoz, 2016). Indeed, digital curation constitutes an increasingly important opportunity for librarians and archivists (Dooley, 2015; Kellam and Thompson, 2016; Munoz and Renear, 2011; Noonan and Chute, 2014; Schaffner and Erway, 2014; Strasser and Krier, 2014; Walters and Skinner, 2011). But libraries' digital curation efforts remain underdeveloped (Tenopir *et al.*, 2014, 2015). Cox and his colleagues (2017) sum up, "extensions of existing library efforts (e.g. training for data literacy, promoting data archives, and rights or IP guidance) may be likely to reach maturity more rapidly than the wholly new services (such as data analysis and visualization)" (p. 2195). Such maturity requires top-down support from library administration and policy as well as bottom-up initiative from information professionals themselves (Tenopir *et al.*, 2015).

Digital curation concerns persist in digital humanities: the DID3 challenge mandated a Data Management Plan describing how the project team planned to curate data generated or collected by the project. PI P53-07 posited, "Curation really is a significant issue; otherwise you're not going to be able to reuse your data yourself, let alone make it available to anyone else."

Yet researchers lack time for digital curation training and tend to learn what skills they have in ad hoc, experiential fashion (Jahnke and Asher, 2012; Tenopir *et al.*, 2012). They want to satisfy preservation requirements in an easy and time-efficient way (Lynch, 2014): as PI P24-01 insisted, "When it comes to managing data and curating data, you could spend an endless amount of time." Unsurprisingly, study participants emphasized the potential for librarians or archivists to curate project data. PI P11-02 mused, "There's a certain moment when you say, 'Shouldn't the library system actually be funding [digital curation] and doing [digital curation]?'"

On one end of the spectrum, participants suggested opportunities for collaboration with librarians and archivists in digital curation. For projects such as DID3, P42-01 asserted:

If you're going to do a big project that's sustainable, that's useful over time, not a boutique one, but something that's supposed to serve a broader audience in a big sustainable way […] you need to look for partners in the libraries and leverage their knowledge and abilities, and for where things are going as well. That's something that is just a lot to ask of a scholar in the humanities, to be up on all those features.

Like P42-01, P4-09 stressed the potential for librarians or archivists to lead preservation planning. Most ambitious, P25-01 reflected:

I wonder if there is a possibility of incorporating archivists and librarians in these projects in that early foundational step, where it's not "oh, here's our project; let's go find an archivist and see how they can help us do what we want to do," but figure out a way where an archivist might be able to come on the team and help ask some of the questions.

Incorporating librarians and archivists into that foundational step would help ensure future fit for purpose. P4-09 asked rhetorically, "What metadata do they need to be able to preserve in the [institutional] Repository […]? Are there things that we're doing incorrectly? Is there any other type of information we should be gathering?"

Such comments point not only to potential collaborations, but also to the importance of taking a lifecycle approach to data to ensure its quality and to provide documentation from the time of its creation, something a DMP should mandate (Borgman, 2009; Ray, 2014). As Schaffner and Erway (2014) comment, "It is one thing to accept a bag of bits, but quite another to ensure that the content will be usable in the future" (p. 13).

On the other end of the spectrum, many participants preferred delegation to collaboration. In these cases the library seemed, in effect, a "dispensary of goods" or a warehouse as opposed to a veritable partner in the research endeavor (Jahnke and Asher, 2012, p. 16). PI P37-04 stated, "All this archiving stuff that we have to do, if that could be taken off my shoulders, that would be really good." Likewise, PI P36-09 claimed, "I want to use the data to study data, and that does not mean saving data and structuring data and things like that."

These participants call to mind Osswald's (2013) observation: "The curation job is hardly glamorous or much admired. For the most part it is a service-oriented back office activity demanding functionality and perfection." Many researchers, whether in DH or other data-intensive fields, characterize such tasks as drudgework, even if crucially important drudgework (Cox *et al.*, 2017).

Despite these participants' interest in a division of labor that would relieve them of a potentially onerous task, collaboration should not be dismissed. As P29-05 shrewdly observed, "It would be lovely if there was just a dedicated department that you could say, "This is all of our research data. Please take care of it." But in reality, it's the same as the "save-my-data-from-the-dying-scientist": they don't know which bits to pick. You have to say, "This is important for this; this is important for this." A lot of preformatting and sorting has to go through before that."

Such sentiments hint at the urgency for librarians to work with researchers at all stages of the research process, for instance in developing DMPs and strategies for data description and preservation writ large (Tenopir *et al.*, 2012, 2014, 2015). Digital curation persists as a remarkable opportunity for librarians and archivists.

ILS education and professional preparation
The three DID3 participants with LIS degrees discussed the imprint of their education on their career trajectories, especially its relevance to their project work. P15-10 entered

graduate school in the late 2000s. She recalled, "When I started the program, the library science track was still very much on cataloging, and a lot of things that I didn't feel like were very data management side of things. I didn't feel like it had yet caught up to what libraries needed to be in the digital information era."

Upon her matriculation P41-10 intended to work as a traditional art librarian. But she "quickly realized that digital libraries was really what was the most interesting thing […] going on within the library world, and it seemed there probably were going to be a lot of job opportunities in that area." As a result, she gravitated toward courses on metadata and information organization. Like P41-10, P28-08 recalled focusing on "less traditional librarianship" in her LIS program. She gravitated toward "a lot of digital, technical kind of areas like databases, web design, all that." "Where it's really come in handy," she concluded, "is the way we think about data and how we can preserve it."

Notwithstanding the content of LIS courses, P41-10 called herself "a big proponent" of getting hands-on experience as preparation for work on projects such as DID3. "When I got to library school," she recalled, "what surprised me was that people were coming in with either very little technical background or skills, or […] with absolutely no experience having worked in a library."

LIS programs might profit from these reflections, especially in the areas of DH and digital curation curriculum development. For instance, LIS students interested in DH grapple with a dual learning curve: technological and domain or disciplinary. Schaffner and Erway (2014) note DH scholars rarely consult the library for domain expertise. Addressing this concern, Senchyne (2016) enjoins engagement with DH "as a thing in and of itself, a demanding thing that requires both/and knowledge" (p. 374). Another scholar advocates integrating DH into required core coursework instead of sidelining it as an elective or as not-for-credit (Moazeni, 2015), and Keener (2015) also exhorts librarians to develop their humanities domain skills and knowledge. These scholars' assertions mesh with those of DID3 participants such as P6-04, who wanted librarians or information professionals who possessed both technical and humanities expertise on their teams. P30-10 expanded upon the skills she thought librarians needed to develop:

> These librarians need a different sort of skills, a good part of which is technology, plus understanding of the scientific disciplines and the digital tools that the librarians have and the publishers have […] They are slowly learning […] but a lot more needs to be done […] in order to offer the services that could be offered by the library, given the digital resources available.

Like DH, digital curation deserves integration into the LIS curriculum for graduates and professionals both: the demand for curation skills and expertise such as that needed in DID3 will only grow (Lynch, 2014; National Research Council of the National Academies, 2015). Acquiring staff with the requisite skill sets remains a serious challenge, as does (re)training existing personnel (Cox *et al.*, 2017; Lyon, 2016; Lyon and Brenner, 2015). Professional education programs such as the Post-Master's Certificate in Data Curation at the University of North Carolina at Chapel Hill can help[2].

Ultimately, curricula must align with current and projected research needs. In particular, programs must offer and encourage students to enroll not only in DH and digital curation courses, but also in courses involving related skills such as text mining, information organization, statistics, information visualization, ontologies, databases, and of course, curation, as this study's participants recommended. Further, programs should encourage, or ideally, require students to accrue practical experience involving technical skills to prepare them for data-intensive professional duties. As yet, the necessary skills and capabilities are not uniformly in place: capacity-building – including recruitment, retraining, and continuing education – takes time (Cox *et al.*, 2017).

Outreach, raising awareness, and organizational culture
On one hand, Keener's (2015) sample of faculty proved "extremely enthusiastic" about and "spoke very highly" of librarians with whom they worked. In this spirit, P32-03 asserted, "The quality of librarians in my lifetime has just increased exponentially." "They have really become very sharp people," she expounded, "good at managing their time, good about thinking through problems."

Respondents such as P15-10 illustrated the resources that librarians could bring to bear on projects such as those in DID3. "We have a lot of tools that people within our institution could use to do a lot of different things. For instance, Altmetric […] could allow them to manage their publications, manage data about their publications […] but a lot of people don't know that we have those tools that are available." P25-01 underlined, "in terms of outreach and making these projects accessible to larger groups, that's also a skill set that I think many librarians and archivists will have, in terms of knowing things like Omeka […] there are so many things that archivists and librarians already know, not just about preserving data and making it safe […] [and] accessible to other people."

On the other hand, some participants discussed the lack of librarians' involvement. Just as many DH scholars fail to appreciate the potential for librarians' involvement (Schaffner and Erway (2014), so too did PI P37-04 confess, "It could be just that there's a lack of communication between my field and the librarians, so I may not know what they can do for me." In much the same way, P39-05 saw a "disconnect between a lot of the library stuff and a lot of the type of stuff that I'm running into." She confessed, "I wouldn't know what to ask for." 24-01 critiqued librarians' transparency, noting, "The things [librarians are] doing internally change quickly and I understand the reasons why, but there's something that seems at times needlessly mysterious about that." Researchers seemed unaware as well of the possibilities for collaboration with LIS programs at their institutions. P28-08 lamented, "We have a library school and no one knows about it."

As these responses indicate, awareness of possible contributions from librarians, whether advisory or technical, seemed low in some participants' responses. In this sense, MSLIS degree holder P28-08 reflected, "That's where LIS feels hobbled to me, the larger problem about the work that we can do because of the stereotypes that fall around librarianship." One of her team's PIs, P12-08 equated LIS work to human resource management: both shared "the paper pusher stigma." As Schaffner and Erway (2014) argue, perceptions and attitudes about the library matter deeply; libraries' organizational cultures must evolve to meet research needs and opportunities, primarily by accepting the imperative of assessment (Lakos and Phipps, 2004).

Responding to such concerns and echoing scholars such as Keener (2015) and Zhang *et al.* (2015), another MSLIS graduate favored librarians "going to these different departments that we could make an impact on, that we have skill sets that could help, and giving a presentation to them or talking to them one-on-one about what we can do for them or the tools that we have to offer, and providing some real-world examples as well, so that it becomes very concrete" (P15-10). Librarian P2-09 chimed in, "We're trying to do more outreach to scholars to be able to do projects like this, so being able to communicate how to exploit an API, or how to exploit data held in a repository, in a way that's not off-putting, would be particularly useful. So there's just a little bit of marketing and/or outreach skilling that's useful."

In terms of outreach, librarians and archivists can start small, as P28-08 suggested: they can provide a "DH-friendly environment," (Schaffner and Erway, 2014, p. 14) tailored to local circumstances (Elliott *et al.*, 2015). More ambitious, in addition to attending department events, they can join pertinent listservs and social media, and serve on institution-wide committees (Vandegrift and Varner, 2013). Data workshops or other opportunities may also prove fruitful, though they must be engineered to avoid being seen by researchers as "just another time suck," to borrow P28-08's phrase.

Digital curation, LIS education, and outreach – these issues demand concentrated attention to maximize the potential return on investment in data-intensive, interdisciplinary, international DH projects.

Conclusions and future research

Stakeholders would do well to consider five recommendations. First, librarians, and archivists' work remains largely invisible in these projects. It must be made visible to exploit existing and to provide evidence for additional resources; a culture of assessment can help. Second, infrastructure must be leveraged and policy developed or clarified for optimal collaboration between librarians and DH. Third, stakeholders need to consider their hiring practices, focusing more on recruiting and retaining those with both technical and domain skills and knowledge. Fourth, digital curation and lifecycle involvement offer librarians and archivists perhaps their most robust opportunity to work with DH scholars. Fifth, LIS educators must align their curricula with current and future research needs and employment opportunities.

Building on this case study, the results of which can only be theoretically generalizable, future research might consider six questions. First, in DH work do librarians themselves prefer to be partners or collaborators or supporters or researchers or some combination thereof? The literature remains equivocal; this study reinforces the ambivalence and ambiguity of librarians' positions in DH work. Notably, despite the optimism evinced in the literature, few participants in this study saw librarians as full partners. P30-10 insisted that librarians "are slowly learning," but stipulated, "a lot more needs to be done."

Second, what specific skills might DH scholars and librarians acquire in order to participate effectively in DH projects? For example, the three LIS-trained participants mentioned domain knowledge, programing and tool development, scientific writing, algorithms, and taxonomies; the practicing librarian mentioned image interaction, data wrangling, and data modeling. What might other librarians involved in DH work suggest in this regard?

Third, what monitoring have funders engaged in with respect to DID3 projects' data management? Requiring a DMP at the application stage is one thing; monitoring changes and compliance in the DMP over the course of the project, much less after, is quite another.

Fourth, wherever and whenever it has been deposited, can data from DID3 projects be reused? Has it been reused and if so, how? P16-06 harkened to this question, observing, "We created this wonderful resource and people want to know how it's being used […] it would be nice to do some sort of survey at some point to find out." Yet project teams may lack incentives for performing such follow-up work, especially once the active grant period elapses. This question bears further scrutiny given the thrust of DID to make publicly funded data reusable.

Fifth, do librarians and archivists help develop the corpora that are used and reused in DH work? If so, how much are they involved and in what ways?

Sixth, how can information professionals such as librarians and archivists lend their skills to project management roles and responsibilities? P28-08's DID3 work stands as a case in point: LIS skills are amenable not only to digital curation, but also to project management. The potential for synergy is clear.

This study supports the DID3 funders' premise that interdisciplinary collaboration in the networked environment nourishes synergies among LIS and DH researchers as they create – or productively fail – to create sustainable knowledge bases more innovative than either discipline could realize alone. Fay and Nyhan (2015) conclude that collaborations may inscribe "a kind of virtuous circle, that benefits all involved […] and this is transferable beyond […] one particular project" (p. 129). To this point, P17-04 explained, "We're […] working with digital preservation people and archivists and trying to integrate the software

into systems that will allow for the preservation […] and the citation of the research data, so I'm excited." Although collaboration takes time to develop and both LIS professionals and academics need relevant training and experience (Fay and Nyhan, 2015; Zhang *et al.*, 2015), initial consultations can engender more robust collaborations over time (Elliott *et al.*, 2015; Keener, 2015). Hence libraries and librarians should involve themselves from the very inception of DH projects (Bryson *et al.*, 2011).

Despite the commonalities between DH and LIS, and their evident if at times subtle congruence as perceived by study participants, the two fields have yet to collaborate in ways that exploit their full potential. Nevertheless, participants praised librarians and libraries, at times effusively: PI P47-08 called librarians "our natural allies" and another PI (P32-03) reflected, "Libraries went from being kind of Sleepy Hollow to being 'where it's at.'" In their research and in their practice – for the two are symbiotic – librarians and archivists must build on this foundation.

Notes

1. https://dev.diggingintodata.org/awards/2009

2. https://sils.unc.edu/programs/post-masters-certificates. Other information science programs that offer continuing education in the area of digital curation include the Digital Information Graduate Certificate (DigIn) at the University of Arizona (https://ischool.arizona.edu/digital-information-graduate-certificate-digin), Dominican University's Certificate in Digital Curation (http://sois.dom.edu/academics/certificates/digital-curation), the University of Maryland's Curation and Management of Digital Assets (CMDA) Post-Master's Certificate (http://cmdacert.umd.edu/), Kent State University's Advanced Study in Digital Preservation Certificate (http://catalog-archive.kent.edu/archive/academics/catalog/2010/collegesprograms/ci/certs/c826.html), San Jose State University's Post-Master's Certificate in Digital Curation (http://ischool.sjsu.edu/programs/post-masters-certificate), and Simmons College's Digital Stewardship Certificate (http://internal.simmons.edu/students/slis/current/courses/curriculum-by-degree/digital-stewardship-certificate).

References

Atkinson, R. and Flint, J. (2004), "Snowball sampling", in Lewis-Beck, M.S., Bryman, A. and Futing Liao, T. (Eds), *Encyclopedia of Social Science Research Methods*, SAGE, Thousand Oaks, CA, pp. 1044-1045.

Bartsherer, T. and Coover, R. (2011), *Switching Codes: Thinking through Digital Technology in the Humanities and Arts*, University of Chicago Press, Chicago, IL.

Bawden, D. and Robinson, L. (2012), *Introduction to Information Science*, Facet, London.

Bazeley, P. (2013), *Qualitative Data Analysis with NVivo*, Sage, Thousand Oaks, CA.

Bell, G., Hey, T. and Szalay, A. (2009), "Beyond the data deluge", *Science*, Vol. 323 No. 5919, pp. 1297-1298.

Bernard, H.R. and Ryan, G.W. (2010), *Analyzing Qualitative Research*, Sage, Thousand Oaks, CA.

Berry, D.M. (2012), *Understanding Digital Humanities*, Palgrave Macmillan, New York, NY.

Borgman, C. (2009), "The digital future is now: a call to action for the humanities", *Digital Humanities Quarterly*, Vol. 3 No. 4, available at: www.digitalhumanities.org/dhq/vol/3/4/000077/000077.html

Brandenburg, M.D., Cordell, S.A., Joque, J., MacEachern, M.P. and Song, J. (2017), "Interdisciplinary collaboration: librarian involvement in grant projects", *College & Research Libraries*, Vol. 78 No. 3, pp. 272-282.

Bryson, T., Posner, M.S., Pierre, A. and Varner, S. (2011), *SPEC Kit 326: Digital Humanities*, Association of Research Libraries, Washington, DC.

Burdick, A., Drucker, J., Lunenfeld, P., Presner, T. and Schnapp, J. (2012), *Digital_Humanities*, MIT Press, Cambridge.

Charmaz, K. (2014), *Constructing Grounded Theory*, Sage, Thousand Oaks, CA.

Choemprayong, S. and Wildemuth, B. (2009), "Case studies", in Wildemuth, B. (Ed.), *Applications of Social Research Methods to Questions in Information and Library Science*, Libraries Unlimited, Westport, pp. 51-61.

Cohen, D.J. and Scheinfeldt, T. (2013), *Hacking the Academy: New Approaches to Scholarship and Teaching from Digital Humanities*, University of Michigan Press, Ann Arbor, MI.

Council on Library and Information Resources (2013), *Research Data Management: Principles, Practices, and Prospects*, Council on Library and Information Resources, Washington, DC.

Courtney, A. and Dallis, D. (2015), "Models for partnering with faculty and supporting digital scholarship", in Sacco, K.L., Richmond, S.S., Parme, S.M. and Wilkes, K.F. (Eds), *Supporting Digital Humanities for Knowledge Acquisition in Modern Libraries*, Information Science Reference, Hershey, PA, pp. 86-106.

Cox, A.M., Kennan, M.A., Lyon, L. and Pinfield, S. (2017), "Developments in research data management in academic libraries: towards an understanding of research data service maturity", *Journal of the Association for Information Science and Technology*, Vol. 68 No. 9, pp. 2182-2200.

Crompton, C., Lane, R.J. and Siemens, R. (2016), *Doing Digital Humanities: Practice, Training, Research*, Routledge, London.

Cunningham, L. (2010), "The librarian as digital humanist: the collaborative role of the research library in digital humanities projects", *Faculty of Information Quarterly*, Vol. 2 No. 2, pp. 1-11.

Dalbello, M. (2011), "A genealogy of digital humanities", *Journal of Documentation*, Vol. 67 No. 3, pp. 480-506.

Deegan, M. and McCarty, W. (2012), *Collaborative Research in the Digital Humanities*, Ashgate, Burlington, VT.

Dooley, J. (2015), *The Archival Advantage: Integrating Archival Expertise into Management of Born-Digital Library Materials*, OCLC Research, Dublin, OH.

Elliott, C., Feeney, M., Kollen, C. and Reyes-Escudero, V. (2015), "A DH state of mind: libraries and the digital humanities", in Sacco, K.L., Richmond, S.S., Parme, S.M. and Wilkes, K.F. (Eds), *Supporting Digital Humanities for Knowledge Acquisition in Modern Libraries*, Information Science Reference, Hershey, PA, pp. 132-155.

Fay, E. and Nyhan, J. (2015), "Webbs on the web: libraries, digital humanities, and collaboration", *Library Review*, Vol. 64 Nos 1-2, pp. 118-134.

Flanders, J. and Munoz, T. (2016), "An introduction to humanities data curation", available at: http://guide.dhcuration.org/intro/

Gardiner, E. and Musto, R.G. (2015), *The Digital Humanities: A Primer for Students and Scholars*, Cambridge University Press, Cambridge.

Gold, A. (2010), *Data Curation and Libraries: Short-Term Developments, Long-Term Prospects*, California Polytechnic State University, San Luis Obispo, CA.

Gold, M. (2012), *Debates in the Digital Humanities*, University of Minnesota Press, Minneapolis, MN.

Gorman, G. and Clayton, P. (2005), *Qualitative Research for the Information Professional*, Facet, London.

Henry, G. (2014), "Data curation for the humanities", in Joyce, M.R. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 347-374.

Hodder, I. (2000), "The interpretation of documents and material culture", in Norman, K.D. and Yvonne, S.L. (Eds), *Handbook of Qualitative Research*, Sage, Thousand Oaks, CA, pp. 703-715.

Jahnke, L. and Asher, A. (2012), *The Problem of Data*, Council on Library and Information Resources, Washington, DC.

Jick, T. (1979), "Mixing qualitative and quantitative methods: triangulation in action", *Administrative Science Quarterly*, Vol. 24 No. 4, pp. 602-611.

Jones, S. (2014), *The Emergence of the Digital Humanities*, Routledge, New York, NY.

Kamada, H. (2010), "Digital humanities: roles for librarians?", *College & Research Libraries News*,
Vol. 71 No. 9, pp. 484-485.

Keener, A. (2015), "The arrival fallacy: collaborative research relationships in the digital humanities",
Digital Humanities Quarterly, Vol. 9 No. 2, available at: www.digi`talhumanities.org/dhq/vol/
9/2/000213/000213.html

Kellam, L. and Thompson, K. (2016), *The Acadeic Data Librarian in Theory and Practice*, ACRL, Chicago, IL.

King, G. (1994), *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton
University Press, Princeton, NJ.

Klein, J.T. (2015), *Interdisciplining Digital Humanities*, University of Michigan Press, Ann Arbor, MI.

Klein, L.F. and Gold, M.K. (2016), *Debates in DH 2016*, University of Minnesota Press, Minneapolis, MN.

Koltay, T. (2016), "Library and information science and the digital humanities: perceived and real
strengths and weaknesses", *Journal of Documentation*, Vol. 72 No. 4, pp. 781-792.

Kretzschmar, W.A. and Gray, W. (2010), "Library collaboration with large digital humanities projects",
Literary and Linguistic Computing, Vol. 25 No. 4, pp. 439-445.

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.

Levine, M. (2012), "Policy, practice, and law", available at: http://guide.dhcuration.org/legal/policy/
(accessed February 1, 2018).

Lincoln, Y.S. and Guba, E. (1985), *Naturalistic Inquiry*, Sage, Newbury Park, CA.

Lynch, C. (2014), "The next generation of challenges in the curation of scholarly data", in Joyce, M.R.
(Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue
University Press, West Lafayette, IN, pp. 395-408.

Lyon, L. (2016), "Librarians in the lab: toward radically re-engineering data curation services at the
research coalface", *New Review of Academic Librarianship*, Vol. 22 No. 4, pp. 391-409.

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.

Mellon, C.A. (1990), *Naturalistic Inquiry for Library Science*, Greenwood Press, New York, NY.

Moazeni, S.L. (2015), "Integrating digital humanities into the library and information science
curriculum", *Public Services Quarterly*, Vol. 11 No. 3, pp. 225-231.

Morgan, D.L. (2008), "Snowball sampling", in Lisa, M.G. (Ed.), *The SAGE Encyclopedia of Qualitative
Research Methods*, SAGE, Thousand Oaks, CA, pp. 816-817.

Munoz, T. (2016), "Recovering a humanist librarianship through digital humanities", in White, J.W. and
Gilbert, H. (Eds), *Laying the Foundation: Digital Humanities in Academic Libraries*, Purdue
University Press, West Lafayette, IN, pp. 3-14.

Munoz, T. and Renear, A. (2011), "Issues in humanities data curation", available at: http://cirss.ischool.
illinois.edu/paloalto/whitepaper/premeeting/ (accessed February 1, 2018).

National Research Council of the National Academies (2015), *Preparing the Workforce for Digital
Curation*, National Academies Press, Washington, DC.

Noonan, D. and Chute, T. (2014), "Data curation and the university archives", *The American Archivist*,
Vol. 77 No. 1, pp. 201-240.

Nyhan, J., Terras, M. and Vanhoutte, E. (2013), *Defining Digital Humanities*, Ashgate, Burlington, VT.

Oliver, G. and Harvey, R. (2016), *Digital Curation*, 2nd ed., Neal-Schuman, Chicago, IL.

Osswald, A. (2013), *Skills for the Future: Educational Opportunities for Digital Curation Professionals*,
European Commission, Florence.

Pannapacker, W. (2012), "Digital humanities triumphant?", in Gold, M.K. (Ed.), *Debates in the Digital
Humanities*, University Press, Minneapolis, MN, pp. 233-234.

Pickard, A. (2013), *Research Methods in Information*, Neal-Schumann, Chicago, IL.

Poole, A.H. (2013), "Now is the future now? The urgency of digital curation in the digital humanities", *Digital Humanities Quarterly*, Vol. 7 No. 2, available at: www.digitalhumanities.org/dhq/vol/7/2/000163/000163.html

Poole, A.H. (2017a), "'A greatly unexplored area': digital curation and innovation in the digital humanities", *Journal of the Association for Information Science and Technology*, Vol. 68 No. 7, pp. 1772-1781.

Poole, A.H. (2017b), "The conceptual ecology of digital humanities", *Journal of Documentation*, Vol. 73 No. 1, pp. 91-122.

Posner, M. (2013), "No half measures: overcoming common challenges to doing digital humanities in the library", *Journal of Library Administration*, Vol. 53 No. 1, pp. 43-52.

Pun, R. (2015), "Conceptualizing the integration of digital humanities in instructional services: possibilities to enhance digital literacy in the 21st century", *Library Hi-Tech*, Vol. 33 No. 1, pp. 134-142.

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.

Robinson, L., Priego, E. and Bawden, D. (2015), *Library and Information Science and Digital Humanities: Two Disciplines, Joint Future?*, Verlag Werner Hulsbusch, Zadaar, pp. 44-54.

Rockenbach, B. (2013), "Introduction", *Journal of Library Administration*, Vol. 53 No. 1, pp. 1-9.

Rubin, H.J. and Rubin, I.S. (2005), *Qualitative Interviewing: The Art of Hearing Data*, Sage, Thousand Oaks, CA.

Sabharwal, A. (2015), *Digital Curation in the Digital Humanities*, Chandos, Waltham, MA.

Sabharwal, A. (2017), "Digital humanities and the emerging framework for digital curation", *College & Undergraduate Libraries*, Vol. 24 Nos 2-4, pp. 1-19.

Sacco, K.L., Richmond, S.S., Parme, S.M. and Wilkes, K.F. (2015), *Supporting Digital Humanities for Knowledge Acquisition in Modern Libraries*, Information Science Research, Hershey, PA.

Saldana, J. (2013), *Coding Manual for Qualitative Research*, SAGE, Thousand Oaks, CA.

Schaffner, J. and Erway, R. (2014), *Does Every Research Library Need a Digital Humanities Center?*, OCLC Research, Dublin, OH.

Schreibman, S., Siemens, R. and Unsworth, J. (2004), *A Companion to the Digital Humanities*, Oxford University Press, Oxford.

Schreibman, S., Siemens, R. and Unsworth, J. (2016), *A New Companion to Digital Humanities*, Wiley Blackwell, Malden, MA.

Senchyne, J. (2016), "Between knowledge and metaknowledge: shifting disciplinary borders in digital humanities and library and information studies", in Klein, L.F. and Gold, M.K. (Eds), *Debates in the Digital Humanities 2016*, University of Minnesota Press, Minneapolis, pp. 368-376.

Sholtz, R.W. and Tietje, O. (2002), *Embedded Case Study Methods*, Sage, Thousand Oaks, CA.

Showers, B. (2012), "Does the library have a role to play in the digital humanities?", available at: https://infteam.jiscinvolve.org/wp/2012/02/23/does-the-library-have-a-role-to-play-in-the-digital-humanities/ (accessed January 11, 2018).

Siemens, R. and Schreibman, S. (2007), *A Companion to Digital Literary Studies*, Blackwell, Malden, MA.

Strasser, C. and Krier, L. (2014), *Data Management for Libraries: A LITA Guide*, American Library Association, Chicago, IL.

Sula, C.A. (2013), "Digital humanities and libraries: a conceptual model", *Journal of Library Administration*, Vol. 53 No. 1, pp. 10-26.

Sutton, B. (1993), "The rationale for qualitative research: a review of principles and theoretical foundations", *Library Quarterly*, Vol. 63 No. 4, pp. 411-430.

Svensson, P. and Goldberg, D.T. (2015), *Between Humanities and the Digital*, MIT Press, Cambridge, MA.

Taylor, S.J. and Bogdan, R. (1998), *Introduction to Qualitative Methods: A Guidebook and Resource*, John Wiley & Sons, New York, NY.

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., Hughes, D. and Allard, S. (2015), "Research data services in academic libraries: data-intensive roles for the future?", *Journal of eScience Librarianship*, Vol. 4 No. 2, pp. 1-21.

Tenopir, C., Sandusky, R.J., Allard, S. and Birch, B. (2014), "Research data management services in academic research libraries and perceptions of librarians", *Library & Information Science Research*, Vol. 36 No. 2, pp. 84-90.

Vandegrift, M. and Varner, S. (2013), "Evolving in common: creating mutually supportive relationships between libraries and the digital humanities", *Journal of Library Administration*, Vol. 53 No. 1, pp. 67-78.

Varner, S. (2015), "Library instruction for digital humanities pedagogy in undergraduate classes", in White, J. and Gilbert, H. (Eds), *Laying the Foundation: Digital Humanities in Academic Libraries*, Purdue University Press, West Lafayette, IN, pp. 205-222.

Vinopal, J. and McCormick, M. (2013), "Supporting digital scholarship in research libraries: scalability and sustainability", *Journal of Library Administration*, Vol. 53 No. 1, pp. 27-42.

Walters, T.O. and Skinner, K. (2011), *New Roles for New Times: Digital Curation for Preservation*, Association of Research Libraries, Washington, DC.

Warwick, C., Terras, M. and Nyhan, J. (2012), *Digital Humanities in Practice*, Facet, London.

Weiss, R. (1994), *Learning from Strangers: The Art and Method of Qualitative Interviewing*, Free Press, New York, NY.

Westbrook, L. (1994), "Qualitative research methods: a review of major stages, data analysis techniques, and quality controls", *Library & Information Science Research*, Vol. 16 No. 3, pp. 241-254.

White, J.W. (2016), "Preface", in White, J.W. and Gilbert, H. (Eds), *Laying the Foundation: Digital Humanities in Academic Libraries*, Purdue University Press, West Lafayette, IN, pp. xi-xiv.

White, J.W. and Gilbert, H. (2016), *Laying the Foundation: Digital Humanities in Academic Libraries*, Purdue University Press, West Lafayette, IN.

Wildemuth, B. (2009), "Existing documents and artifacts as data", in Wildemuth, B. (Ed.), *Applications of Social Research Methods to Questions in Information and Library Science*, Libraries Unlimited, Westport, pp. 158-165.

Williford, C. and Henry, C. (2012), *One Culture: Computationally Intensive Research in the Humanities and Social Sciences, A Report on the Experiences of First Respondents to the Digging into Data Challenge*, Council on Library and Information Resources, Washington, DC.

Yin, R. (2009), *Case Study Research: Design and Methods*, Sage, Los Angeles, CA.

Zhang, Y., Liu, S. and Mathews, E. (2015), "Convergence of digital humanities and digital libraries", *Library Management*, Vol. 36 Nos 4-5, pp. 362-377.

Further reading

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.

Benardou, A., Constantopoulos, P., Dallas, C. and Gavrilis, D. (2010), "Understanding the information requirements of arts and humanities scholars", *International Journal of Digital Curation*, Vol. 5 No. 1, pp. 18-33.

Borgman, C. (2015), *Big Data, Little Data, No Data*, MIT Press, Cambridge.

Kapur, M. (2008), "Productive failure", *Cognition & Instruction*, Vol. 26 No. 3, pp. 379-425.

National Endowment for the Humanities (2014), "NEH announces the winners of the 2013 digging into data challenge", available at: www.neh.gov/news/press-release/2014–01-15 (accessed June 1, 2017).

Palmer, C.L., Teffeau, 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.

Riley, J. (2014), "Metadata services", in Joyce, M.R. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 149-165.

Smith, M. (2014), "Data governance: where technology and policy collide", in Joyce, M.R. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 45-59.

Westra, B. (2014), "Developing data management services for researchers at the university of oregon", in Joyce, M.R. (Ed.), *Research Data Management: Practical Strategies for Information Professionals*, Purdue University Press, West Lafayette, IN, pp. 375-391.

Corresponding author
Alex H. Poole can be contacted at: poole.alex@gmail.com

Appendix

DID round	Projects funded	Countries involved	Number of funding agencies involved	Monies disbursed
DID1	Using Zotero and TAPOR on the Old Bailey Proceedings: Data Mining with Criminal Intent (DMCI)	USA; UK; Canada	4	$2 million
	Digging into the Enlightenment: Mapping the Republic of Letters			
	Towards dynamic variorum editions (DVE)			
	Mining a year of speech			
	Harvesting speech data sets from the Web			
	Structural analysis of large amounts of music information (SALAMI)			
	Digging into Image Data to Answer Authorship Related Questions (DID-ARQ)			
	Railroads and the Making of Modern America			
DID2	An Epidemiology of Information: Data Mining the 1918 Influenza Pandemic	Canada; Netherlands; UK; United States	8	$4.8 million
	Cascades, Islands, or Streams? Time, Topic, and Scholarly Activities in Humanities and Social Science Research			
	Chartex			
	Digging by Debating			
	Digging Into Connected Repositories (DiggiCORE)			
	Digging Into Human Rights Violations: Anaphora Resolution and Emergent Witnesses			
	Digging Into Metadata: Enhancing Social Science and Humanities Research			
	Electronic Locator of Vertical Interval Successions (ELVIS): The First Large Data-Driven Research Project on Musical Style			
	Imagery Lenses for Visualizing Text Corpora			
	IMPACT Radiological Mummy Database			
	Integrated Social History Environment for Research (ISHER)-Digging Into Social Unrest			
	Integrating Data Mining and Data Management Technologies for Scholarly Inquiry			
	Mining Microdata: Economic Opportunity and Spatial Mobility In Britain, Canada and the USA, 1850-1911			
	Trading consequences			
DID3	Automating Data Extraction from Chinese Texts	Canada; Netherlands; UK; United States	10	$5.1 million

(continued)

Table AI.
Digging into Data
funded projects

DID round	Projects funded	Countries involved	Number of funding agencies involved	Monies disbursed
	Commonplace Cultures: Mining Shared Passages in the 18th Century Using Sequence Alignment and Visual Analytics			
	Cleaning, Organizing, and Uniting Linguistic Databases (COULD)			
	Digging Archeology Data: Image Search And Markup (DADAISM)			
	Project Arclight: Analytics for the Study of Twentieth-century Media			
	Global Currents: Cultures of Literary Networks, 1050-1900			
	Mining biodiversity			
	Digging into Linked Parliamentary Data (DiLiPaD)			
	Field Mapping: An Archival Protocol for Social Science Research Findings			
	Digging into Signs: Developing Standard Annotation Practices for Cross-Linguistic Quantitative Analysis of Sign Language Data			
	Trees and tweets: mining billions to understand human migration and regional linguistic variation			
	Legal structures			
	Mining relationships among variables in large data sets from complex systems (MIRACLE)			
	Resurrecting Early Christian Lives: Digging In Papyri In A Digital Age			

Table AI.