CASE STUDIES

Advancing Digital Repository Services for Faculty Primary Research Assets: An Exploratory Study

Stephen Kutay * 
California State University Northridge, 18111 Nordhoff St., Northridge, CA 91330, USA

A R T I C L E   I N F O

Article history:
Received 23 May 2014
Accepted 28 August 2014
Available online 27 October 2014

Keywords:
Digital curation
Digital repositories
Digital services
Library collaboration
Preservation
Scholarly communication

A B S T R A C T

The Oviatt Library at California State University Northridge (CSUN) hosts two digital repositories represented by Digital Collections for archival and historical materials, and ScholarWorks institutional repository (IR) for scholarly output. This paper reports on an exploratory study for advancing digital repository services regarding faculty primary research assets created in the course of research and/or collected by scholar custodians of archival materials at CSUN. A survey was distributed to understand: 1) which faculty and departments collect or create primary source assets as part of their research, 2) what types of assets are collected or created, 3) the activities performed to preserve these assets, 4) the level of interest in making primary research documents available online, 5) faculty knowledge of library methods, and 6) attitudes regarding collaboration with the library. This survey functions as part of a needs assessment toward the development of new and enhanced digital repository services to advance research, preservation, data curation, instruction, and exhibition. This knowledge will also help to systematize library and faculty collaboration through the development of policies and workflows that reduce ad hoc re-evaluations and protracted negotiations over the ability of the library to support digital research and instruction projects.

* 25662 Chimney Rock Rd., Valencia, CA 91355, USA. Tel.: +1 818 677 2335.
E-mail address: stephen.kutay@csun.edu.

I N T R O D U C T I O N

The Oviatt Library at California State University Northridge (CSUN) hosts two digital repositories represented by Digital Collections for archival and historical materials, and ScholarWorks institutional repository (IR) for scholarly output. They are further differentiated by resource origins and types, document lifecycles, metadata entry, and methods of arrangement. However, some redundancies exist. Digital Collections contain examples of faculty and student works as well as secondary sources. ScholarWorks preserves and disseminates primary sources as raw data, and over time becomes populated with documents at the end of their lifecycles. Despite well-defined domains, neither repository can wholly support every type of resource or need associated with its use. Recent inquiries into leveraging these dual systems to support more general asset management obscure the boundaries of their original intent. Meanwhile, issues in scholarship have emerged, which these fixed systems do not comprehensively address. These issues include functionality to support bulk resource downloads into lecture presentations (Johnston, 2011); asset management, retrieval and preservation across distributed systems (Freeman, Ward, Jones, Molloy and Snow, 2010; Marshall, Bly and Brun-Cottan, 2007); resource aggregation and communication support for networked research teams (Borgman, 2007) and; resistance to data sharing (Borgman, 2012). Our digital services development must logically factor extensibility and scalability to accommodate a larger array of resources, and, more importantly, the myriad ways they are used.

Advancements in flexible open source asset management systems are extending digital services. The Hydra Project, for example, is a collective effort among academic institutions across the United States to develop service enhancements for digital assets, institutional repositories, digital media libraries, archives, and workflow management (Hydra, n.d.). The project presents a forum in which partner institutions address local needs with the benefit of sharing their experience across the network. Initial results are promising and may ultimately offer other institutions viable, open source alternatives for internal collections and asset management. This model demonstrates that there is a reciprocal influence between expressed needs and technological capabilities. As a result, digital services require a perpetual state of evaluation.

Services assessment is inherently a local process intended to reveal discontinuities between resources and stakeholder needs. Reflective analyses of local library infrastructures and funding sources enable the deployment of new services. Stakeholder needs also vary from institution to institution, department-to-department and scholar-to-scholar. Therefore an academic services assessment should consider local factors such as institutional objectives, disciplinary foci, and embedded research behaviors to determine the course of strategic development. It is because of this variation that a local analysis is detrimental to the success of new services.
This paper reports on an exploratory study to advance digital repository services regarding faculty primary research assets at CSUN. The term ‘assets’ is preferred here to describe information resources owned or created by scholars (or by association, colleges and departments) and used in the course of developing research output, instructional materials or to promote educational activities. A primary research asset is thus defined as any primary source document used to support research and teaching, which may or may not include raw data and data sets. In addition to investigating assets that are created, this study examines those that are collected by scholar custodians of archival materials used for the purpose of research and instruction.

A survey of the faculty at CSUN was distributed to understand: 1) which faculty and departments collect or create primary source assets as part of their research, 2) what types of assets are collected or created, 3) the activities performed to preserve these assets, 4) the level of interest in making primary research documents available online, 5) faculty knowledge of library methods, and 6) attitudes regarding collaboration with the library. This survey functions as part of a needs assessment toward the development and support of new and existing digital repository services to advance research, preservation, data curation, instruction, and exhibition. This knowledge will also help to systematize library and faculty collaboration through the development of policies and workflows that reduce ad hoc re-evaluations and protracted negotiations over the viability of digital research and instruction projects.

LITERATURE REVIEW

The literature selected in the following review examines issues directly relevant to the academic digital repositories at CSUN, but also investigates related issues of information management and sharing from both the personal and organizational perspectives. In addition, a library services framework is also considered to help guide future planning as a result of this study.

In 2003, Clifford Lynch of the Coalition for Networked Information (CNI) envisioned fundamental changes in digital libraries regarding new forms of scholarly communication, interoperable systems, diverse community needs, and flexible solutions to address the surge of personal collections of documents (Lynch, 2003). The re-imagining of library systems tested the preconceived boundaries of repositories to manage access to locally published materials alone (Conway, 2008). A 2007 census of IR activities in the United States showed a significant increase in media formats and document types such as ETDs, images, learning objects and software (Markey, Soo, St. Jean, Kim and Yakel, 2007). The range of scholarly works in combination with needs for repurposing them introduced new tensions for repositories to function as broader digital asset management systems (Conway, 2008; Waters, 2006). Repositories that aggregate diverse resources are especially valuable as a tools for research and instruction (Johnston, 2011), and facilitate the ways in which faculty use research materials (Conway, 2008; Davis and Connolly, 2007).

More recently, faculty research is supported through dedicated library services and embedded partnerships to manage ingest, description, access and preservation of research data (Brown and Tucker, 2013; Carlson, Ramsey and Kotterman, 2010). Though data management services are now administered across numerous academic institutions and required for NSF and NIH funding (National Science Foundation, 2011; National Institute of Health, 2003) the efficacy of sharing research data is challenged by disciplinary norms, ethics, lack of incentives (Borgman, 2012) and complex, proprietary formats requiring extensive context to interpret (Borgman, 2012; Weber, Baker, Thomer, Chao and Palmer, 2012). For Science, Technology, Engineering and Math (STEM) scholars, the perceived importance of sharing data is high (Borgman, 2012; Scaramozzino, Ramírez and McGaughey, 2012). Yet scholars in many disciplines are reluctant to share their data (Borgman, 2012), and those who do, have difficulties with retrieving and deciphering data among collaborators (Freiman, Ward, Jones, Molloy and Snow, 2010; Scaramozzino, Ramírez and McGaughey, 2012).

The role of research libraries as the de facto leaders in data curation has been called into question. Data are often “domain-specific” products generated by scientific organizations outside of the academy that require domain-specific knowledge and technologies to properly organize, interpret and preserve them (Nielsen and Hjørland, 2014; Weber, Baker, Thomer, Chao and Palmer, 2012). It is suggested that academic librarians embark on analytical studies of their own institutional domains (Nielsen and Hjørland, 2014) in order to best serve their constituencies.

From the archival perspective, digital repositories have played a significant role in how students and scholars are using records in their research. The need to re-conceptualize archival materials as long-term assets has a value that “resides in the ability to repurpose them for different uses, audiences, and situations.” (Yakel, 2004). This view of records as assets is profoundly impacted by their use in the digital domain.

Archival materials used in education have never been more promising, in part, due to the availability of digital resources for eLearning. Still, fundamental barriers persist. In a study of the use of primary sources for teaching History, for example, researchers have found them to be essential. However, these resources are most often selected from published works, rather than from online or physical archives (Malkmus, 2010). Teachers are insufficiently aware of digital collections available online (Malkmus, 2010). Despite reported improvements in student engagement using digital primary sources, some teachers lacked the experience to efficiently navigate across various digital information systems (Diekema, Leary, Haderlie and Walters, 2011). Given such obstacles, some research suggests that digital primary sources are promising for areas outside the historical frame, such as teaching critical analysis. Students using digital archival materials for problem-based learning (PBL) have demonstrated higher levels of performance and satisfaction (Chen and Chen, 2010).

The notion of archives, and by extension digital collections, as instruments of social justice and community agency has received greater attention as a focus of the archival profession (Bastian, 2013; Caswell, 2013; Duff, Flinn, Suurttam and Wallace, 2013; Gilliland, 2011; Harris, 2002a; Jimerson, 2009; Ketelaar, 2001). Digital and digitized records online provide a public space with which to engage such issues as LGBTQ identity (Wakimoto, Bruce and Partridge, 2013), post-colonial transition (Harris, 2002b; Mckemnith, Faulkhead and Russell, 2011) and living archives (Rhodes, 2014). Under these and other contexts, will Social Science faculty leverage the library’s archives and digital repositories to disseminate evidence generated from the communities they research?

Researchers of Personal Information Management (PIM) have taken a pragmatic, user-centered approach to understanding patterns of behavior regarding the storage, retrieval and preservation of research information. Scholarly authors view the assets used in their research and publications as essential to their scholarly output, and consider these to be component parts of the published product, deserving some level of archival treatment (Marshall, 2008). In a pilot study, Marshall, Bly and Brun-Cottan (2007) discovered four specific themes with regard to personal archiving challenges. These themes included 1) issues with appraising one's own materials, 2) highly distributed storage environments, 3) curatorial issues for large aggregations of files, and 4) the lack of long-term access support in desktop systems (Marshall, Bly and Brun-Cottan, 2007). Data preservation is a concern of scholars, however a lack of documentation and insufficient knowledge of appropriate preservation formats impedes progress in this area (Freiman, Ward, Jones, Molloy and Snow, 2010). Through collaborative and consultative services, archivists and librarians are well positioned to observe these issues and convey strategies for faculty information management.