



Sustaining Data Archives over Time: Lessons from the Organizational Studies Literature

Kalpana Shankar & Kristin R. Eschenfelder

To cite this article: Kalpana Shankar & Kristin R. Eschenfelder (2015) Sustaining Data Archives over Time: Lessons from the Organizational Studies Literature, New Review of Information Networking, 20:1-2, 248-254, DOI: [10.1080/13614576.2015.1111699](https://doi.org/10.1080/13614576.2015.1111699)

To link to this article: <http://dx.doi.org/10.1080/13614576.2015.1111699>



Published online: 17 Dec 2015.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Sustaining Data Archives over Time: Lessons from the Organizational Studies Literature

KALPANA SHANKAR

*School of Information and Communication Studies, University College Dublin, Belfield,
Dublin, Ireland*

KRISTIN R. ESCHENFELDER

*School of Library and Information Studies, University of Wisconsin-Madison, Madison,
Wisconsin, USA*

This article will explore how organizational theoretic approaches to understanding institutional sustainability could be applied to contemporary data archives and archival practices, with the aim of raising larger questions about data sustainability. The article is part of a larger study in progress of how one particular example of long-lived archives, Social Science Data Archives (SSDAs), have maintained themselves over time. SSDAs predate both computers and the Internet and provide a unique opportunity to examine perceptions about what makes an archive sustainable, through the ups and downs of funding cycles and massive changes in technical and organizational infrastructure.

KEYWORDS *social science data archives, sustainability, business models, organizational theory*

INTRODUCTION

In some respects, digital libraries and data repositories are integral to furthering research in almost all disciplines. They are increasingly promoted as useful to communities well beyond the academy, including policy makers, educators, and civil society. However, their ongoing maintenance has become a pressing concern for practitioners and groups and individuals

© Kalpana Shankar and Kristin R. Eschenfelder

Address correspondence to Kalpana Shankar, School of Information and Communication Studies, University College Dublin, Belfield 4, Dublin, Ireland. E-mail: Kalpana.shankar@ucd.ie

who rely on them for their work. Without assurances that the institutions and practices that curate data will last over time and will do so in ways that maintain the integrity and usefulness of the data, data depositors and users will not trust in the longevity of data, with significant negative implications for research and practice.

With some exceptions, institutional longevity has not received much attention vis-à-vis repositories, cyberinfrastructure/e-science, or digital libraries until fairly recently (Chowdhury 2013; Rieger and Warner 2010). Chowdhury defines sustainability as having three key components: economic, social, and environmental. Digital libraries, he argues, must be affordable to access, ensure equitable access to information, and should do so while implementing environmental best practices.

We extend this frame to suggest another dimension of sustainability that increasingly needs attention: that of the institution. Although Chowdhury's three components are arguably significant for institutional sustainability of data organizations over time, we argue that institutional/organizational longevity subsumes and transcends these other components. We need to pay specific attention to how organizations maintain themselves; not just economically, but as part of the larger institutional networks in which they figure, such as universities, professional organizations, and other archives (Eschenfelder, Shankar, Downey, and Lin 2015).

In this article, we briefly examine three interrelated questions:

- How do social science data archives (SSDA) in particular contribute to the conversation on sustainability?
- How do the theoretical and practitioner literatures on archives and organizations frame the topic of sustainability?
- What do the gaps in the conversations suggest for new work?

For this article, we focus on how three components of sustainability are articulated and enacted: the role of business models, institutional relationships, and the value of the services provided. The goal of this article is not to be a systematic review or even a substantive content analysis of the relevant literatures, but instead a preliminary attempt to coalesce disparate literatures and ideas for insight into the nature of data sustainability and to articulate some themes and ideas for further investigation.

SOCIAL SCIENCE DATA ARCHIVES (SSDA) AS EXEMPLARS OF LONGEVITY

This article is drawn in part from our larger study on the history of social science data archives (SSDA); preliminary empirical results have been presented

elsewhere (Eschenfelder et al. 2015; Shankar et al. 2014). SSDA are exemplars of data longevity, in part because they are the most long-lived examples of data archives, some dating back 50 to 60 years without a break in activity or service. SSDA predate the Internet, indeed the widespread prevalence of computers. Some of the earliest SSDAs still exist today, having adapted themselves and their services to new data sources, users, technologies, and cycles of institutional change. It has been argued that the fairly narrow kinds of data sets that have traditionally been in the purview of SSDA limit the kinds of researchers, technologies, and delivery mechanisms that form the ecosystem of SSDA with the result that SSDA are “easier” to maintain over time than more complex data institutions that vary more widely in their scope and scale (Corti 2012). Our research has instead found that the kinds of relationships that SSDA have created with other institutions, users, and technologies can provide a lens into understanding other contemporary data institutions think about their own persistence.

FRAMING SUSTAINABILITY

How do institutions adapt? The concept of “resilience” has increasingly gained prominence; that is, the capacity of a system to adapt to change and continue to develop. Resilience requires balancing the need for stability and order with an institution’s need to adapt to disruption and change. Burnard and Bhamra (2011) refer to resilience as “the emergent property of organizational systems that relates to the inherent and adaptive qualities and capabilities that enable an organization’s adaptive capacity during turbulent periods” (5587). Developing and maintaining information infrastructures require interrelated levels of labor to create systems that function over time: the institutionalization of generating goods and services that users want and value, the enacting technologies that deliver such goods and services, and robust organizational institutional relationships that support the technical infrastructure of information delivery (Ribes and Finholt 2009). For data archives, the institutional “work” consists of adapting business models, maintaining networks of relationships, and articulating the value and need for data to depositors, users, and funders.

Business Models

Although SSDAs and other data archives encompass and incorporate numerous business models (sometimes within the same institution, usually evolving over time), both the organizational and practitioner literature articulate the need for business models that are in accord with the archive’s mission and yet can weather disruption.

Generally, two interrelated missions are invoked for SSDAs in general and data archives in particular. SSDAs were created in the 1960s to make the abundance of quantitative data from the longitudinal surveys of political scientists, government agencies, census bureaus, and nongovernmental organizations accessible and available to social science researchers (Gilliland-Swetland 1996) through the technologies available at hand to do so. This central mission has not changed (although the disciplines, kinds of data, and quantities of data have of course evolved over time). Business models have adapted to partner with universities and government and funding agencies, balance “freemium access” with fee structures and revenues for data deposit (Brophy 2006; Maron and Loy 2011) and maintenance, and adopt and develop standards and technologies appropriate to scope and scale of the archive (Hamilton 2004). In short, business models for archives are adapted to the larger ecosystems of local, national, and even international institutional politics. To give one example, data archives must be able to assure researchers and data depositors of secure data for ten or twenty years (for example) but this agreement can be bolstered or threatened by shifting national and host institution priorities and concerns.

Multi-Institutional Relationships

Sustainability of organizations is contingent upon the networks in which they operate. The introduction of new tools for data distribution, preservation, description, and dissemination spawn the ability to create new spatial and temporal relations. In turn, new opportunities for the actors to coordinate and sustain collaboration and coordination come into play (Ribes, Jackson, Geiger, Burton, and Finholt 2013). For the SSDA, this would argue that data archives need to be conceptualized as part of a larger network of archives, universities, libraries, data depositories, and standards bodies (Baker 2007; Lefurgy 2009). In a practical sense, this takes form in multiple ways: securing grants and contracts with data depositors, creating relationships with host institutions such as universities, forging multi-institutional partnerships for data sharing (to spread the workload and secure both breadth and depth of coverage) and outreach, and collaborating with researchers. The web of relationships takes numerous forms. For example, since the 1960s, SSDAs have offered workshops and classes, both at disciplinary meetings and on locations, to introduce researchers to their offerings and train them in their use. In turn, researchers become institutional champions, advocating for the licensing of services for their own libraries and universities. In other cases, SSDAs have been at the forefront of data standards work and enrolling new archives. Collaborations may also increase revenue streams as has been pointed out in the digital libraries literature (Maron and Loy 2011); smaller institutions do not need to “reinvent the wheel” but instead can participate in

the development and maintenance of a wide set of technical infrastructures that scale up in terms of efficiency and delivery.

The Value of Data Services

How do stakeholders advocate for the services they need? As revenues, incomes, institutional relationships evolve, evaluation, and renewal are regularly revised. Furthermore, the justification of intellectual merit and broader impacts must be continuously attuned to the disciplines data archives support. Decreased funding for archives or changing institutional relationships throw open questions of relevance. Open access, open data, regulatory mandates (from funding agencies) and disciplinary norms to publish data alongside peer-reviewed publications constitute disruptions that challenge the underlying mission of SSDAs: do we still need data archives if “anyone” can publish data to the Web? Or do we need them even more because “anyone” can publish data to the Web—and SSDAs provide a trusted entity that will insure longevity and integrity? In some cases, the disciplinary training of investigators varies over time. Some specializations become less relevant and other specializations must be added to make new objects researchable. SSDAs, for example, responded to the trend for labor and social history and political science in the 1970s by acquiring data sets of new interest to researchers. In all of these contexts, SSDAs have had to make their case and their mission clear and enroll the support of their stakeholders. This requires clear focus on the nature of the archive’s mission, commitment to stakeholders, and assessment of the resources needed to continue providing value as users see it (Pojasek 2013; Rieger and Warner 2010). Strengthening existing collaborations and the development of new partnerships signals value to the larger network of archives.

SUMMARY

Significant public and private investments into data archives and infrastructures, the reliance of researchers and institutions on their availability, evolving regulatory and disciplinary norms with respect to data sharing, all point to a growing conversation on making data sustainable. Rather than emphasizing the technical preservation of data, a widespread discussion that has yielded numerous fruitful developments, this article takes a brief look at one component of that conversation that still seems undertheorized and discussed (and more intractable in some ways): making the organizations that manage data sustainable over time and adaptable to change.

SSDA give us a long-standing lens into the problem of data sustainability: they have thrived for almost 60 years alongside their flashier “cyberinfrastructure” or “e-science” counterparts, many of whom have folded without

a long-term commitment to making their data available to their constituent communities. The practitioner literature of digital libraries and data archives has focused on the specific and pragmatic goals of economic sustainability: fee structures, specifics of contracts, and resource management. Infrastructure studies have focused on human and technological practices to align scientific infrastructures with constituent communities and stakeholders. Lastly, organizational studies and theory, which have looked at how institutions function as systems and as parts of networks, give us a larger framework in which to slot these conversations. All emphasize the need for robust business models that reflect the mission of the institution (even as that mission evolves), the web of relationships in which the data institution participates, and articulation of the value of data and services being provided.

We suggest that an urgent gap in the discussion is the implications of dissonant timescales for evaluation, and in turn, articulating value. How does an archive know it is “successful”? What data is “valuable”? Income, data sets archived/downloaded, and citations constitute some tangible and measurable outcomes that can be reported to decision-makers, but how measures are selected remain open topics for discussion (and necessity). Such measures can clash with funding cycles, technological obsolescence, and institutional change: all of which operate on different timescales. We expect that the scholarly and practitioner conversations in data preservation will increasingly look at these issues across particular data/digital archives and institutions to develop answers and new questions.

ACKNOWLEDGMENTS

The authors would like to acknowledge Rebecca Lin and Jennifer Nygren McBurney for their assistance on the project.

FUNDING

The authors would like to thank the Alfred P. Sloan Foundation, the Irish Research Council, the Wisconsin Alumni Research Foundation, and ASIS&T History and Foundations SIG History Fund for financial support.

REFERENCES

- Baker, K. S. “New Opportunities for Research Libraries in Digital Information and Knowledge Management: Challenges for the Mid-sized Research Library.” *J. Libr. Adm.* 46.1 (2007) 65–74. Print. doi:<http://dx.doi.org/10.1300/J111v46n01>

- Brophy, P. "Projects into Services: The UK Experience." *Ariadne* 46 (2006). Web. 5 Nov. 2015. doi:http://dx.doi.org/10.1300/J111v46n01_05
- Burnard, K., and R. Bhamra. "Organisational Resilience: Development of a Conceptual Framework for Organisational Responses." *Int. J. Prod. Res.* 49.18 (September 2011): 5581–99. Web. 5 Nov. 2015. doi:<http://dx.doi.org/10.1080/00207543.2011.563827>
- Chowdhury, G. G. "Sustainability of Digital Libraries: A Conceptual Model." In *Research and Advanced Technology for Digital Libraries*, Aalberg, T., Papatheodorou, C., Dobrev, M., Tsakonas, G., Farrugia, C. J. (Eds.). Springer Berlin Heidelberg, 2013, pp. 1–12. Print.
- Corti, L. "Recent Developments in Archiving Social Research." *Int. J. Soc. Res.* 15.4 (2012): 281–90. Web. 5 Nov. 2015. doi:<http://dx.doi.org/10.1080/13645579.2012.688310>.
- Eschenfelder, K. R., K. Shankar, G. Downey, and R. Lin. "Sustainability of Social Science Data Archives: A Historical Network Perspective." *IASSIST 2015*, June 3, 2015, St. Paul, Minn, 2015. Print.
- Gilliland-Swetland, A. "Social Science Data Archives in the New World?" *For the Record: Data Archives, Electronic Records, Access to Information and the Needs of the Research Community*. Ed. Rena Lohan, Mark Conrad, Ken Hannigan, & John A. Jackson. Dublin, Ireland: Institute of Public Administration, 1996. 54–63. Print.
- Hamilton, V. "Sustainability for Digital Libraries." *Libr. Rev.* 53.8 (2004): 392–95. Web. 5 Nov. 2015. doi:<http://dx.doi.org/10.1108/00242530410556210>
- Lefurgy, W. G. "NDIIPP Partner Perspectives on Economic Sustainability." *NDIIPP Partner Perspectives on Economic Sustainability* 57.3 (2009): 413–26. Print. doi:<http://dx.doi.org/10.1353/lib.0.0044>
- Maron, N. L., and M. Loy. *Revenue, Recession, Reliance: Revisiting the SCA/Ithaka S+R Case Studies in Sustainability*, 2011. Web. 5 Nov. 2015. <<http://www.sr.ithaka.org/publications/funding-for-sustainability-how-funders-practices-influence-the-future-of-digital-resources/>>.
- Pojasek, R. B. "Organizations and Their Contexts: Where Risk Management Meets Sustainability Performance." *Environ. Qual. Manag.* 22.3 (2013): 81–93. Web. 5 Nov. 2015. doi:<http://dx.doi.org/http://10.1002/tqem.21338>
- Ribes, D., and T. A. Finholt. "The Long Now of Technology Infrastructure: Articulating Tensions in Development." *J. Assoc. Informat. Syst.* 10.5 (2009): 375–98. Web. 5 Nov. 2015. <<http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1507&context=jais>>.
- Ribes, D., S. Jackson, S. Geiger, M. Burton, and T. Finholt. "Artifacts that Organize: Delegation in the Distributed Organization." *Info. & Org.* 23.1 (2013): 1–14. Web. 5 Nov. 2015. <<http://www.sciencedirect.com/science/article/pii/S1471772712000371>>.
- Rieger, O. Y., and S. Warner. "Developing sustainability strategies for arXiv. Inf." *Stand. Q.* 22.4 (2010): 17–18. Print.
- Shankar, K., K. Eschenfelder, G. Downey, P. Fleming, C. Engerson, R. Lin, and J. McBurney. "Social Science Data Archives: Case Studies in Data Sustainability." *iConference 2014*, Berlin, Germany, March 11–15, 2014. Web. 5 Nov. 2015. <<http://hdl.handle.net/2142/47387>>.