

Building Relationships “A Foundation for Digital Archives”

Jonathan Crabtree¹, Darrell Donakowski²

The Odum Institute UNC Chapel Hill 22 Manning Hall Chapel Hill N.C.¹

Jonathan_Crabtree@unc.edu

ICPSR University of Michigan P.O. Box 1248 Ann Arbor MI²

dwdonako@UMich.edu

Abstract. Now that archives have entered the digital age, what is the best way to build and sustain them? Collaborating with other archives could improve how they operate and grow, but building relationships consumes time and scarce resources. Is it worth the effort? To test this concept, the Odum Institute, ICPSR and four major digital archives are collaborating through the Library of Congress’s National Digital Information Infrastructure and Preservation Program. Their partnership is called Data-PASS. The technical infrastructure to support digital archives is constantly changing. Digital archivists must find solutions by building relationships with developers. Our partnership’s efforts to develop a federated approach to data archives require building relationships at the producer, administrative, and program application development levels. The institute is collaborating with Harvard-MIT Data Center to adopt its Virtual Data Center as a platform for its archive. While our focus is social science data, the approach would work in many fields.

Keywords: Management, Design, Human Factors, Standards, User issues.

1 Introduction

Throughout the world universities, businesses, governments, and archives are facing the task of building as well as maintaining digital archives. Archive administrators must navigate increasingly complex and technical waters to accomplish this task. They must make decisions on metadata standards, database standards, and interface standards. IT managers and archivist must choose between a growing number of innovative open source options and the security and support traditionally provided by more expensive proprietary systems [1]. A common thread among these options is the relationships needed to be successful. Regardless of the technical solution chosen, the need for collaborative relationships is constant.

2 Data-PASS Experiences

The Data Preservation Alliance for the Social Sciences (Data-PASS) was created with an award from the U. S. Library of Congress’s National Digital Information Infrastructure and Preservation Program (NDIIPP) [2]. Data-PASS is led by the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan. Other partners include the Roper Center for Public Opinion Research at the University of Connecticut, the Howard W. Odum Institute at UNC Chapel Hill, the

Henry A. Murray Research Archive, a member of the Institute for Quantitative Social Science at Harvard University, the National Archives and Records Administration, and the Harvard-MIT Data Center, also a member of the Institute for Quantitative Social Science at Harvard University. The partnership's primary goal is to identify and preserve historically-significant digital social science data at risk of being lost.

Building relationships has been critical to success, but time consuming. Although relationships come in many shapes and sizes, the central ingredient necessary for success is trust. Principal investigators must trust that archives will value and preserve their research. Archivists must trust software developers to deliver reliable, innovative platforms for storing, retrieving, and analyzing important digital data.

2.1 Developing Relationships

From the beginning, founding members of Data-PASS have had a good working relationship. We share similar goals and missions, and this was an incentive to work together on this project. Despite the competitive academic world each partner lives in, they have been able to find common ground by contributing to the effort in areas of specialty and strength.

Relationships developed with your researchers, i.e., data producers, are critically important. Building and maintaining those relationships are the most important tasks an archivist has.

Strong relationships often begin with education. Researchers need to understand the benefits of archiving data. During this educational process it is important to remember your audience. Some are researchers from academic institutions, but private non-profit research organizations also produce vast amounts of digital data. Often they are willing to share and archive those data, but need help overcoming organizational barriers.

2.2 Interaction with Software Developers

Exploring different archival software packages leads down many paths. Custom building a system might require collaborating with a team of developers. Sometimes the most affordable option is to adopt technology developed at other institutions. Regardless of the approach, these collaborations call for strong communication between developers and archive administrators. For example, the Odum Institute is testing the Virtual Data Center (VDC) as an open source platform to upgrade our archive, and we have been working closely with its developers, the Harvard-MIT Data Center [3]. Although the Odum Institute has one of the oldest and largest archives of machine readable data in the U.S, it is important to keep adapting to a constantly changing IT environment [4]. Another area of collaboration Data-PASS is exploring is the ability to federate with other partners. Under a federated approach, users will be able to search and download data from all partnering archives, as well as run advanced statistical calculations.

2.3 Adoption of Standards

Adopting common standards for information systems lays the ground work for relationships to grow and prosper. Common metadata standards among partners are essential, and the social sciences have chosen DDI as a common metadata standard. DDI provides the in depth tagging structures needed for complex datasets and is compatible with the Dublin Core [5]. The Harvard-MIT development team used the Open Archive Interface (OAI) to allow the federation of archive information between partners [6]. Using common technical standards has allowed partners that expose their metadata using an OAI server to join Data-PASS and keep technical costs within acceptable limits.

3. Conclusions

The Data-PASS partnership has allowed its members to do many things that would not have been possible alone. From upgrading technical infrastructure to adopting the same metadata standards, partners have benefited far beyond their investment of time and energy. Digital archives are, by definition, connected to the world in a way never before possible. New developments in software and hardware present challenges, but also opportunities for building and strengthening digital archives. In the future, those who succeed will likely find success through collaboration and through building enduring relationships.

Acknowledgments We extend our thanks to the Odum Institute and ICPSR as well as our Data-PASS partners.

References

1. Altman, M., 2003. "Open Source Software in Libraries: from Greenstone to the Virtual Data Center, and Beyond", *IASSIST Quarterly* 25(1):5-12
2. Data-PASS. *Web site*. <URL: <http://www.icpsr.umich.edu/DATAPASS/>> (04/26/06)
3. The Data.Org Project. *Web site*. <URL: <http://thedata.org/index.php/Main/AboutVDC> (04/26/06)
4. The Odum Institute. *Web site*. <URL: <http://www.odum.unc.edu>> (04/26/06)
5. The DDI Initiative. *Web site*. <URL: <http://www.icpsr.umich.edu/DDI/codebook/index.html> (04/26/06)
6. OAI Project. *Web site*. <URL: <http://www.openarchives.org/> (04/26/06)