

EDITORIAL BOARD

Editor and Chair

Mary Beth Weber, *Rutgers University*

Members

Julian Everett Allgood, *New York University*

David Banush, *Brown University*

Sian Dorian Brannon, *University of North Texas*

Barry Brown, *University of Montana*

Karen E. K. Brown, *State University of New York-Albany*

Steven B. Carrico, *University of Florida*

Christopher J. Cronin, *University of Chicago*

Nadine P. Ellero, *Auburn University*

Lisa B. German, *Penn State University*

Andrew S. Hart, *University of North Carolina at Chapel Hill*

Steven A. Knowlton, *University of Memphis*

Philip E. Schreur, *Stanford University*

Michele Seikel, *Oklahoma State University, Stillwater*

Anne M. Sleeman, *Community College of Baltimore County, Catonsville*

Lori Terrill, *University of Wyoming*

Lynn N. Wiley, *University of Illinois at Urbana-Champaign*

Virginia Kay Williams, *Wichita State University*

Oksana Zavalina, *University of North Texas*

Ex-Officio Members

Norm Medeiros, *Haverford College*
Book Review Editor, *LRTS*

Alice Pearman, *YBP*
Editor, *ALCTS News*

Charles Wilt, Executive Director,
ALCTS

Guest Editorial

The Digital Preservation Imperative: An Ecosystem View

Brian E. C. Schottlaender

Brian E. C. Schottlaender is the Audrey Geisel University Librarian at the University of California, San Diego. I am pleased to have a guest editorial from a former ALCTS president to provide through *LRTS* a larger audience for this important and emerging topic. This guest editorial is adapted from an essay that appeared in the (UC San Diego Library) *Faculty File*, Spring 2013, pages 2–3.

As the information universe becomes increasingly digital, there is a growing need to preserve digital assets that represent the intellectual capital of scientific disciplines, educational communities, and government and cultural agencies. This need is both quantitative and qualitative in nature. Digital resources, particularly digital data, are proliferating at a staggering rate. According to the International Data Corporation (IDC), the amount of data worldwide grew 48 percent between 2011 and 2012 to 2.7 zettabytes, or 2.7 billion terabytes.¹ Additionally, digital resources are qualitatively different from analog resources (print and media) in terms of fragility and complexity.

Digital information resources are fragile in ways that differ from analog information resources, largely because they are far more dynamic. Consider the following:

- they are easily and frequently revised/updated, linearly (v. 1.0, v. 2.0, v. 3.0, etc.) or cumulatively
- they may be available in various “views” (e.g., a data set rendered in SQL looks very different from the same data set rendered in Visual Studio)
- they can be more easily altered by someone other than the original creator
- they are more susceptible to corruption over time
- the storage media on which they reside typically have a far shorter life span than their analog storage counterparts

However passé it may be, paper, for the most part, is pretty durable.

The most immediate and significant consequence of the dynamic nature of digital information resources is that their preservation calls for a much more active process than that required for analog resources. Passive preservation (“put it someplace cold and dark and throw away the key”) simply will not work in the digital environment. The bits have to be kept moving and need to be checked and rechecked to ensure that they do not become compromised or succumb to data decay.

Digital resources are not just more fragile than their analog counterparts—they are also more complex. In the analog world, a book appears to be a wonderfully simple thing. Scan it into digital form, however, and it becomes a “complex digital object,” full of individual elements (i.e., pages) that must relate to each other in a certain order, an order that must be preserved if the book is to be

readable. Moreover, it is easy to link from one digital object to another, creating an even more complex digital object that raises questions about what exactly should be preserved. Some types of resources (multimedia, for example) are completely dependent on the software that renders them usable, yet others, such as e-books, are also dependent on the hardware required to make them accessible.

While preservation has never been a single-agency undertaking, this combination of prolificacy, fragility, and complexity calls for an ecosystem approach to digital preservation, and to digital stewardship, in general. This approach includes three essential elements: access, management, and preservation. Curators tend to view this ecosystem as a cycle, whereas technologists see it more as a stack. Regardless of how one views it, the components are by and large the same.

Following are examples of each ecosystem element:

- The access component is manifest in portals like the Pacific Rim Library (PRL), developed by the Pacific Rim Digital Library Alliance (<http://prl.lib.hku.hk/exhibits/show/prdla/browse-collections>); Calisphere, developed by the University of California's (UC) California Digital Library (www.calisphere.universityofcalifornia.edu); and the Digital Public Library of America (DPLA), developed by a coalition of libraries led by Harvard (<http://dp.la>).
- The management component is exemplified by DSpace, open source repository software developed by the Massachusetts Institute of Technology; Fedora (Flexible Extensible Digital Object Repository Architecture), a digital asset management architecture developed by the University of Virginia (UVA); and the Digital Asset Management System (DAMS), developed by the UC San Diego Library (UCSD).
- The preservation component is well represented in Chronopolis, developed by UCSD, the University of Maryland, and the National Center for Atmospheric Research (NCAR); HathiTrust, developed by UC and the University of Michigan (UM); and the Academic Preservation Trust, under development at the University of Virginia.

These ecosystem elements have multiple and variable relationships with one another. Some of the UCSD content managed in DAMS is syndicated for discovery purposes in Calisphere and replicated for preservation purposes in Chronopolis, for example.

The newest player to emerge in the ecosystem is the Digital Preservation Network (DPN) led by UV, Stanford University, the University of California, UM, and the University of Texas. DPN (pronounced "deepen") was conceived as a backbone to unite and provide common services to the preservation elements of the ecosystem, including services like transmission, replication, auditing, and succession. Similar to Internet2, moreover, DPN is conceived as being of, by, and for the academy. As such, it is a direct response to the "growing need to preserve digital assets that represent the intellectual capital of scientific disciplines [and] educational communities" that served as the point of departure for this editorial.

Reference

1. Frank Gens, "Top 10 Predictions: IDC Predictions 2012: Competing for 2020," accessed April 29, 2013, <http://cdn.idc.com/research/Predictions12/Main/downloads/IDCTOP10Predictions2012.pdf>.