

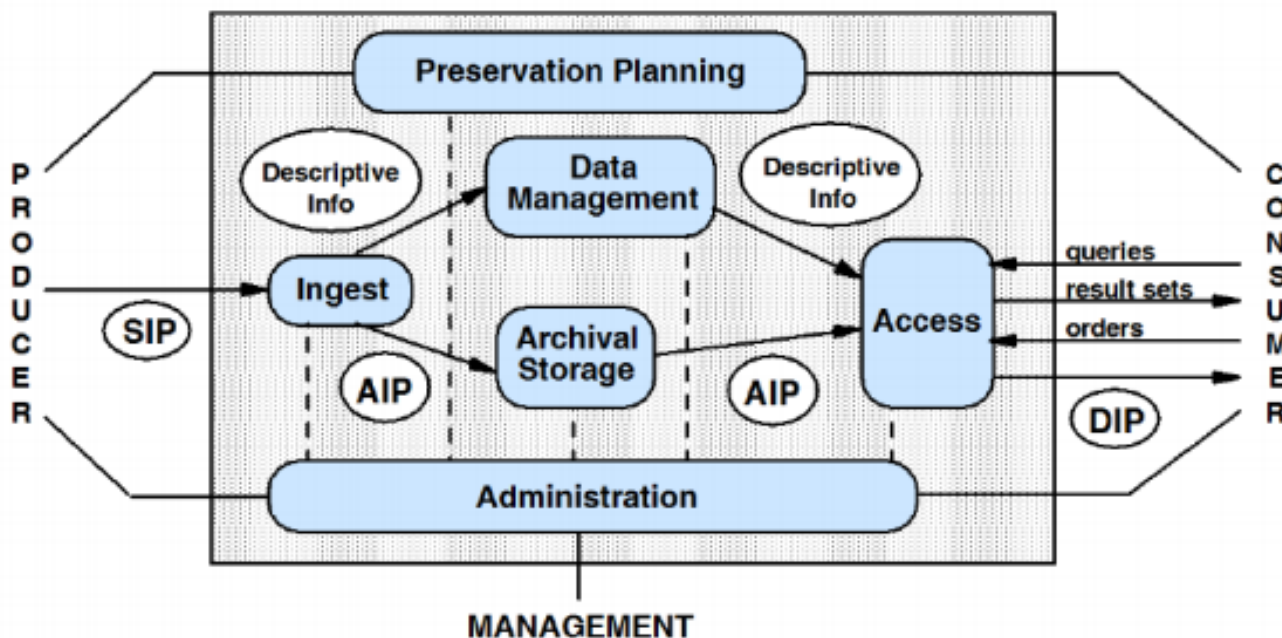
Digital Preservation: How much is it going to cost, and who can I work with?

Jody L. DeRidder | September 21, 2016 | Reynolds Journalism Institute

The expense of digital preservation for the news producer will vary depending on how much of the effort is managed in-house. By collaborating with those who already have the infrastructure, the cost to news agencies could be very little indeed. For example, news publishers in Kentucky were already regularly submitting their PDFs to a vendor for management of legal notices.

The University of Kentucky Special Collections obtained permission to electronically collect this news content from the vendor. It now has taken over all the preservation processes that follow, which includes embargo of access to protect the publishers from income loss. The University of North Texas regularly sends hard drives to publishers, onto which they load PDF eprints, and return them by mail.

The Florida Digital Newspaper Library and Caribbean Newspaper Digital Library found that the most productive method to transfer news content was to harvest the files on a regular basis from the electronic edition subscription services already in use. In this case, most publishers provided a subscription to the Digital Library System. The first and sometimes only cost for publishers is developing a



Digital preservation is complex but manageable. All it takes is a commitment to managing the process. Above is a diagram overview of the process from "Reference Model for an Open Archival Information System (OAIS)," 2012.

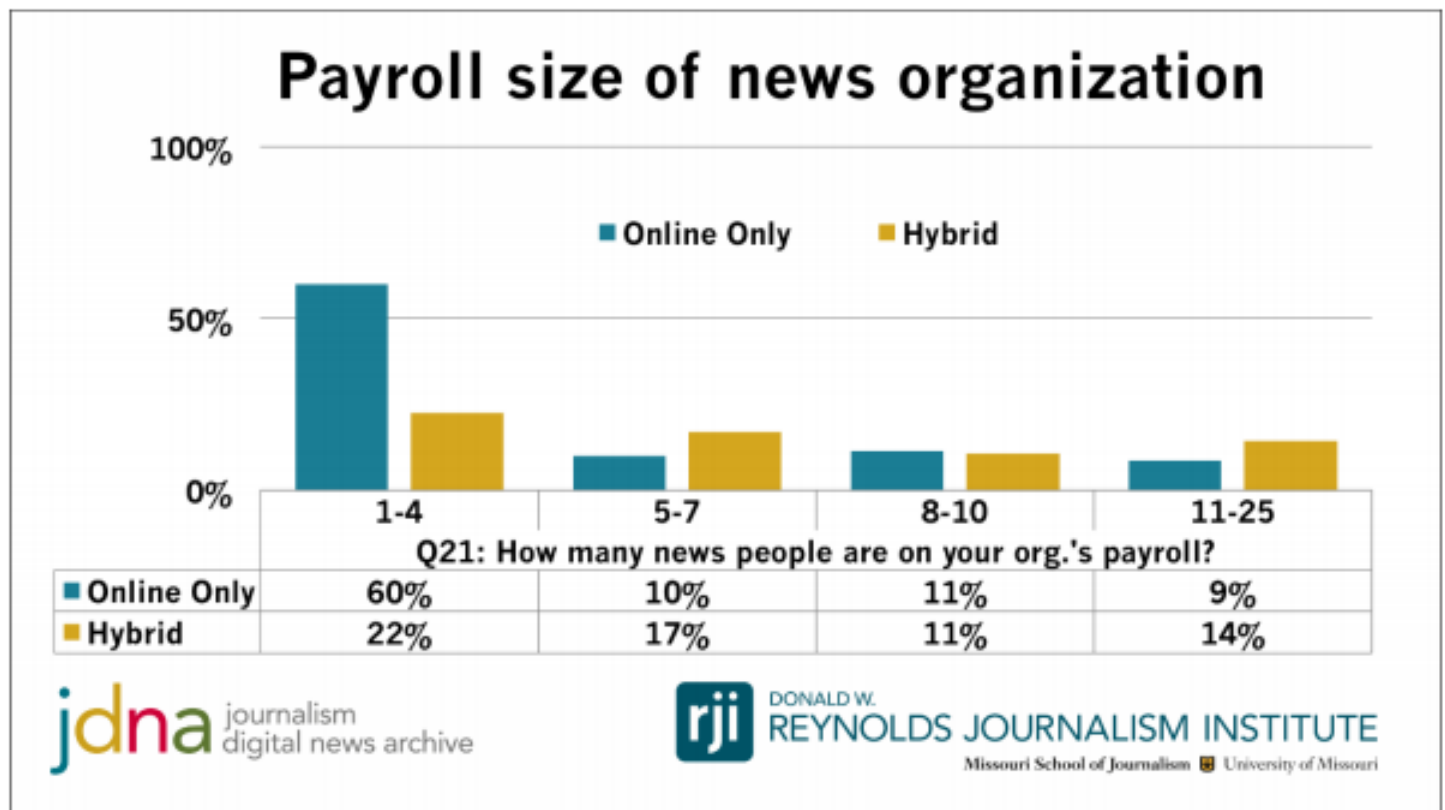
method for preservationists to access the content, providing rights for them to work with the material and providing access. This will only work if appropriate supporting agencies can be located and collaborations established.

Complete costs for digital preservation are difficult to establish. The functional model for preservation used widely today was developed by the Management Council of the Consultative Committee for Space Data Systems (CCSDS) and is called the Open Archival Information System (OAIS). In this model the steps for content management include ingest into an archival system, storage and management of the content and then access. Success of the model depends on proper planning and administration. Digital preservation is complex but manageable. All it takes is a commitment to managing the process. Above is a diagram overview of the process from "Reference Model for an Open Archival Information System (OAIS)," 2012.

Clearly, the institutions that manage the majority of this process also bear the brunt of the costs. The challenges to cost modeling have been organized into those of the business application, cost information, technology and methodology. For the business of preserving news materials, cost information will vary based on how much of what types of content need to be preserved and the methods used for preparation, storage, management and access. Technology is a moving target; no one can know what tomorrow's hardware and software will require, so how can we know the cost of migrating today's and yesterday's news so it can be used tomorrow?

Estimating costs

Still, a number of approaches are available to estimate costs. The Royal Danish Library and Danish National Archives provide an online spreadsheet editable in Google Docs called "Cost Model for Digital Preservation". In their examples the send and receive functions are excluded, and they



Results from a 2014 phone survey indicating the number of news people on payroll at online or online and print news organizations.

approximate costs by incorporating varying wage levels for different tasks levels: manager, computer scientist and technician. This tool expresses costs in terms of monetary costs and/or person weeks, and storage time is expected to be 25 years. Within that migration of content is needed approximately every 8 years. Storage costs alone can be affected by numerous factors, as described by David Rosenthal of Stanford University.

Clearly, the process of digital preservation is not one to be undertaken lightly. How much of this effort is it realistic for news publishers to undertake?

Consider the steps involved:

- Inventory and tracking of content across all media
- Format normalization, monitoring and migration
- Servers dedicated to manage content
- Duplication of content geographically, and backup systems
- Development of standardized metadata: structural, descriptive, administrative, technical and rights
- Access provision, preferably from a single interface

Digital Preservation is a lot of work! Yet many news agencies have very few personnel. Sixty percent of respondents in a 2014 survey conducted by the Donald W. Reynolds Journalism Institute said their online news agencies have only 1 to 4 news people on their payroll, and only 6 percent had a news librarian. The amount of staff time available for digital preservation is minimal at best.

Potential partners

But news agencies aren't the only ones who care about preserving the record. At the 2011 Newspaper Archive Summit, stakeholders included historians, genealogists, and local communities. Government agencies and memory organizations across the U.S. have demonstrated their concern for digital preservation of news since 1982 in the United States Newspaper Program and the National Digital Newspaper Program. These programs were underwritten by the National Endowment for the Humanities and the Library of Congress.

As more and more memory institutions develop the architecture and establish similar processes, the task becomes that of making the necessary connections between publishers, memory organizations and funders.

Potential stakeholders include press associations, newspaper archives, newspaper producers, journalism schools, journalists and researchers.

Others who might be interested include the Internet Archive, The Associated Press, the International Press Telecommunications Council and the Newspaper Association of America.

Vendors such as Newz Group could play an important role. News representatives from major organizations and media companies need to be involved. Universities, research libraries and historical organizations will likely provide the infrastructure and expertise for preservation.



Success stories

Several states already have developed programs to assist publishers in preserving born-digital news as well. In addition to the Texas Digital Newspaper Program, the Kentucky PaperVault Project, the Digital Library of the Caribbean, and the Florida Digital Newspaper Program, other notable projects include the Minnesota Digital Newspaper Project and the Born Digital Project of the California Digital Newspaper Collection. These and others have laid the groundwork for how to preserve news long term.

With the recognition that more can be achieved in collaboration than alone, Katherine Skinner at the Educopia Institute has proposed the development of a "National Action Assembly" to explore the issues involved, train advocates and catalyze connections between stakeholder communities. In an effort to engage top minds on this topic in the fields of journalism, library science, information technology, law, government and philanthropy, the Reynolds Journalism Institute is hosting a conference October 13 -14, 2016 entitled "Dodging the Memory Hole: Saving Online News".

Now is the time to locate funding agencies, stakeholders and memory organizations to assist news publishers in preserving the recent and current record of history.