

## Problem Solving Pioneers Plot Library of the Future

By Tara Caimi

A version of this article first published in [Penn State's IT Stream Magazine](#)

~

For centuries, libraries have existed to collect, preserve, and ensure access to information, and unprecedented advances in technology over the past several decades have done nothing to change this mission, according to Mike Furlough, associate dean of research and scholarly communications for the Penn State University Libraries. The methods and models employed to accomplish those goals, however, have radically shifted.

In the early 1990s, several developments combined to create a perfect storm of information technology. Multiplying computational power and the ability to share research and information across the globe via the Internet set the stage for an ensuing explosion of information and data sharing potential that would shape the future of research.

The Human Genome Project exemplifies the outcome of this combination. Slated for completion in fifteen years, the project culminated in just thirteen due to vast technological advancements and, according to Furlough, the fact that “everyone who was working on it had agreed to share the results of their work publicly and create a big data bank of the genome.”

As computational power increased, so did the amount of data generated. “What took three years to decode the first time can be done in a matter of hours or days now in labs with a sample of DNA,” says Furlough. “The amount of genetic information that’s been created through funded research in the past decade is greater than the federal government thinks they can actually keep up with,” he adds.

### **Information overload**

Today’s libraries face related challenges with increasing urgency. “Federal funding agencies have become very concerned about long-term access to data and research that they’re funding,” says Furlough, who believes libraries are in a unique position to contribute to solutions with regard to creating the services, tools, and spaces that help researchers store, document, and share data.

Digital collections curator Patricia Hswe co-chairs a Research Data Management team at Penn State, providing guidance and assistance to faculty and administrators on developing plans to manage data generated by their research efforts. Though these data management services came as a direct result of new requirements exclusive to those submitting research proposals to the National Science Foundation, Hswe believes this type of planning is critical for all research in the digital age.

Important up-front considerations revolve around information access and preservation, two long-term specialty areas of libraries that have been widely affected by digital formats. “Our emphasis is on service and resources,” says Hswe, adding that the goals of developing data management plans include enabling faculty to conduct research more efficiently and to reach a wider audience with their results.

Years of experience have taught Furlough that collaborating with researchers at the planning stage leads to the best preservation and organization strategies for long-term access. He began his library career at the University of Virginia as an English graduate student in the mid-1990s. As he explored the potential of the web to advance teaching and learning and to facilitate access to data, he worked on a variety of digitizing projects and became more involved with faculty and graduate students who were using available technologies in novel ways.

Several projects piqued Furlough’s interest in research, but his work with a religious studies professor examining the Salem Witch Trials stood out from the rest. The professor investigated geographic locations of the accused, among other aspects of demography with regard to those involved. “By using this technology, Furlough says, “he was able to come up with new theses that actually contradict the previously accepted theories about the dynamics of the village at that time.”

Six years into his position at Penn State, Furlough smiles when remembering his time as a graduate student. “I ran away to the library,” he says, “and I never looked back.”

### **Penn State and the digital age**

While Furlough toiled in the research enterprise at UVA, Sue Kellerman, Judith O. Sieg Chair for Preservation, launched what would become a long-term digitizing program at Penn State. Prior to a phone call she received from the dean of libraries in 1992, Kellerman focused solely on preserving paper collections. When the dean told her of the “brand new tool for digitizing collections and making them available online,” Kellerman recalls,

nobody knew how the technology would affect libraries. Penn State was one of eight institutions at the time to use the new scanning technologies in their libraries.

Though digitizing collections represented a new concept in the early 1990s, technology has enabled research for the past fifty years, according to Furlough. “The earliest GIS (geographic information systems) were created in the 60s,” he says, “primarily to help landscapers, but they were quickly being adopted into other fields during that time.”

Alluding to the intrinsic role of technology in library work, Linda Friend, head of Scholarly Communications Services at the University Libraries, notes that libraries have been creating metadata in the form of cataloguing “for a hundred years.” While computer programmers popularized the term metadata in the 1990s for purposes of their work, “it’s really just information fields,” Friend says. “It’s the title and the author. A catalogue record is metadata.”

At Penn State, technology is central to the enterprise of scholarly communications, in which Furlough sees his role quite clearly. “I help the library develop programs that help researchers and students exploit technologies to disseminate and preserve their work,” he says, adding that the broader field of scholarly communications encompasses the entire system of distributing, accessing, and preserving research. Outlets can include peer-reviewed publications as well as those less formal such as blogs or conference presentations. Libraries, according to Furlough, are uniquely positioned to examine changes in research practices and explore ways in which the system of scholarly communications may better support research and discovery.

Illustrating the impact of scholarly communications, Kellerman points to three volumes of a civil war-era diary digitized about a year ago. Since those volumes were digitized and made available online, several new projects, including transcriptions and text about the author of the diary have been produced.

Scholars from various fields including genealogy and history have shown interest, as have war reenactors, according to Kellerman. “This sat in the vaults since the 1860s. No one knew it was there. Now we have it digitized, and everyone is jumping all over this. It’s producing new scholarship.”

## **Evolving research methods and services**

In libraries, the practice of scholarly communications once focused on

digitizing efforts such as purchasing materials, scanning, and posting materials online, but current trends more closely examine the library's role in better serving faculty and students at the individual level. "We need to broaden our conversations and really ask about the ways we can deploy our technology and our expertise to serve the interests of the students and researchers," Furlough says.

To that end, Hswe and Friend consult with researchers to learn about their needs and offer guidance from the planning stage to future access options. With a Ph.D. in Russian language and literature, Hswe began her library experience as a researcher. Her background in Russian literature led her to the Slavic Library at the University of Illinois, where she was offered a postdoctoral fellowship and where her work "behind the scenes" in the library graced her with a fresh perspective.

"As a researcher I did not perceive librarians enough as a resource," Hswe says, "so I'm really committed to trying to foster this kind of understanding about librarians and libraries as resources that should be mined more by faculty and students."

Hswe appreciates advances in technology that allow for better tracking of the ways in which people are using online data. The ability to see which pages people most often visit, the links they click on those pages, and the files they access fosters more informed decisions with regard to the types of information and services people currently use and will be likely to need in the future.

In addition to creating and implementing usability testing, providing information on copyright and intellectual property laws, optimizing materials for search engines, and running analytics on digitized collections, Hswe works directly with researchers, as does Friend. "We're really trying to figure out what kinds of things we can turn into actual services that we can put out on the web and say this is what the library can do for you," says Friend, who admits to once balking at the idea of becoming a librarian.

A voracious reader as a child, Friend remembers her mother suggesting that she pursue library studies in college. "I was like, are you kidding me? I would never do that." Her attitude changed when, as a teenager, she took a position in a public library. Here, the value of service provided through libraries became clear, and her mother's intuition proved prognostic when she chose to pursue a higher degree in library studies.

Since starting her career as a reference librarian at Penn State in 1978, Friend's role has changed both dramatically and continually. She attributes

her longevity in the position to the changing nature of the work. “I was sure I’d be waltzing all over, and I would live near the ocean, for sure,” she says, adding that the word “librarian” does not effectively describe the many roles it entails. Friend has performed everything from computer programming, to service development, to project management, and advancements in technology over the past several decades have added new complexities to those roles.

### **Preservation from print to digital**

Some challenges faced by today’s libraries involve developing the best ways to select, organize, and preserve online publications. Processes to preserve print collections, including everything from rebinding and deacidification techniques to tracking and auditing, have been developed and perfected over hundreds of years. Finding ways to preserve digital materials with the same level of quality is a current concern.

Sitting in her office chair, Kellerman reaches toward one of the many stacks of books and papers that line the floor in front of the wall-to-wall bookshelf. She picks up a thick book, the approximate size of a bible, with a faded black hardcover barely clinging to the warped and yellowed pages within. “This is a brittle book,” she says. “I used to be so concerned about books, about paper. Actually, this is becoming less and less a concern because the new digital media is more fragile.”

While the physical evidence of book and paper deterioration is visible, the fragility of digital materials manifests in less obvious ways. Servers may crash, for example, and files can corrupt. Also posing a prominent threat is the likelihood that storage formats will become obsolete (e.g. floppy disks).

Kellerman looks at these digital format concerns in relation to her roots. With a passion for preservation she attributes, in part, to her grandmother, Kellerman sought work at the University Libraries to save money for graduate school in preparation for a career in museums. Her work at the library inspired her to change directions, and she instead pursued a master’s degree in library science at the University of Pittsburgh.

Kellerman later returned to Penn State to carry out a project funded by the National Endowment for the Humanities in which she was tasked with locating, identifying, and cataloguing newspapers in Pennsylvania. As she traversed the state, visiting publishers’ offices, museums, libraries, and private collectors’ homes in search of newspapers to catalogue, the question of preserving newspapers often arose. “So that’s where the intersection of my passion for preservation and libraries joined,”

Kellerman says, but it was a disaster at Pattee Library that solidified her career in preservation.

Through working at the University Libraries during a time when people still smoked in the stacks and ate food while handling collections, Kellerman, along with several colleagues, recognized the need for a preservation program. In 1993, fate would grant them their wish. “We had a water main break outside of Pattee Library and water flooded the bottom floor stacks. That disaster had a silver lining,” Kellerman remembers, “and I was made full-time preservation officer.”

After more than two decades of working with print materials, Kellerman considers today’s digital formats and the future they will necessitate. “Storing, preserving, curating, the whole life cycle of the digital asset—that’s the future that preservationists will be working in,” she says as she rolls her chair across the office to the overstuffed bookshelf and pulls out a pristine paperback titled *Pennsylvania Deer and their Horns*. The original publication date is 1915. “We have this project with the Penn State Press,” she says. “These are books that are coming out of our stacks that we’re going to digitize and republish.”

### **Technology—the unknown variable**

While online publishing models provide greater opportunities for discovery of articles, research, and information, it is clear that new models come equipped with new sets of hurdles, preservation notwithstanding. Contrary to common conceptions of open access, publishers have imposed greater fees and restrictions on the use of online materials. Rather than purchasing and owning materials, libraries now often rent digital data sets at the mercy of each publisher’s unique use and access restrictions.

Friend uses the term “born digital” to describe such data and has witnessed the results of data sets purchased from commercial vendors whose access requirements were not supported by available technology. Rather than being accessible to a broad audience, the files remained on a hard drive to be used by one person at a time, voiding the open access potential of the digital format.

Another consideration prompted by technological advances, as evidenced in the Human Genome Project, is the potential for research to generate unwieldy amounts of data. Libraries have yet to determine their roles in capturing, organizing, preserving, and sharing that data. “It’s a real challenge to know how to serve a discipline like the life sciences now as it’s evolving very rapidly, and what their needs will be over the next ten years,

we don't know," says Furlough.

Other disciplines face similar questions, as evolving formats, such as audio and video, come with growing or atypical storage requirements. A current effort at Penn State seeks to address such challenges through strategic collaboration. The University Libraries and Digital Library Technologies (DLT), a unit of Information Technology Services have teamed up to create a digital repository for the many products of Penn State research.

Building the digital repository rather than purchasing a proprietary system is intended to ensure long-term agility, to meet the unique needs of Penn State researchers, to implement Penn State-specific policies and standards, and to limit dependence on an external system that is subject to unexpected change. The Libraries' challenge, according to Furlough, lies in determining the level and scope of service that can be provided and maintained over many years.

### **Libraries of the future**

"It can seem overwhelming," Furlough admits. When coaxed into predicting the future, he projects a shift toward greater institutional collaboration with regard to providing the level of support and services necessary to serve the developing needs of researchers and universities. "I think libraries and even IT organizations in universities are going to find that they cannot support the research enterprise within their institutions by themselves the way they've done in the past. They're going to have to work in tandem with each other on campus but also with other IT organizations at other schools or with other libraries at other schools," he says.

Kellerman predicts shifts in the business model of future libraries to include the elimination of duplicates and more focus on materials and collections targeted to specific curricula and communities of users. "How many Time magazines do we really need?" she says, acknowledging that the processes needed to carry out such a transition have yet to be addressed.

Several years ago, as the Penn State Digitization and Preservation unit prepared to embark on a major scanning effort, Kellerman learned of Google's efforts to do the same. "I just had this vision of a big harvester machine coming down to every library and picking up everything they could scan," she says. To avoid redundancy, she made the choice to concentrate on digitizing special collections and rare materials.

Since that time, other university libraries have shifted their digitization efforts in the same direction, as special collections provide unique value in each university library. “This will help to ensure libraries’ relevance in the future,” says Kellerman, noting that trust is an emerging concern with regard to relying on companies such as Google for mass digitization of materials.

While materials digitized through the University Libraries are managed, tracked, and stored under controlled measures, Google’s potential as a trusted repository remains unclear. “Are they storing files in a way that in fifty years we can access them?” Kellerman questions. Should I say if it’s digitized in Google, it’s done for the good of mankind? I don’t know that,” she says, indicating that future libraries must consider preservation of digital content not only for their own constituents, but also for humanity at large. “Collectively,” Kellerman says, “this is the library’s responsibility.”

### **Problem solving pioneers**

The field of scholarly communications is changing, according to Furlough, but “it isn’t quite clear how it’s changing.” Further, he says “It’s not clear what’s going to be on the other side of all this change.”

Hswe notes vast differences in today’s research methods and tools than those used when she was involved in the field. “We have a dynamic set of users in faculty and students who, themselves, are always changing, and the tools that they use are always changing.” As supporters of that research, libraries are finding ways to adapt.

Kellerman addresses changes in technology by hiring for new skills in the area of digital preservation. “Every new person we hire has to bring in new skills,” she says, confident that a new generation of preservationists will find solutions to ensure that digital materials are effectively preserved.

At the same time, Friend and Hswe conduct their consultations with faculty and graduate researchers to identify services that will be of highest long-term value. Rather than fearing the unknown quality of future issues, these problem solving pioneers embrace the opportunity to think differently about the tools and services libraries have to offer.

In contemplating collections, an enduring staple of libraries, Hswe admits that her work in the digital realm has altered her perception. “It’s like a brave new world toying with the idea of a collection in an online environment,” she says, considering the application of traditional collection development practices to digitized content. “Are we at a point

where it makes more sense to strategize for content rather than collections?”

With more questions on the horizon than answers, the path for libraries of the future is unclear. Furlough believes new technologies and changes in the field have allowed libraries to ask new questions and provide services in different ways. “The changes represent more of an evolution than a revolution,” he says.

At core, librarians are still concerned with organizing information, collecting and preserving data, determining how to make that data accessible to people in ways that are most useful, and guiding users to information that is reliable and valid. “That whole issue of validity of information is still going to be there, and it may keep librarians in business,” says Friend.

Not only are librarians in a unique position to guide users to the best resources and to point them to reliable and authentic information, according to Hswe, but librarians can also assist with complex copyright and intellectual property concerns that have surfaced in response to reproducing materials in digital format.

Libraries have always had to be flexible and to generalize, Furlough points out, qualities he feels place libraries in a positive position to respond to new challenges and to change in general. “We thrive on change,” Kellerman says of library professionals, and her colleagues are eager to agree.

“Working on solutions is the positive—this is the exciting part of the job,” says Hswe. “I think it’s probably one of the most stimulating professions at this point because there’s so much going on and so many opportunities.”

Traversing the nebulous space between the library of the past and the library of the future, today’s library synthesizes with the information technology storm triggered two decades ago that shows no signs of abatement. Despite the uncertainties, Friend sees more prospects than challenges. “We’ve got tremendous resources with the things we can do with technology,” she says, “but the field will be ambiguous for a long time to come.”

In considering future roles with respect to this ambiguity, one certainty emerges to the delight of these library professionals. Their jobs, it is clear, will never be boring.