The Preservation of Electronic Government Information (PEGI) Project is a two-year initiative aimed at addressing national concerns around the collection and preservation of born-digital government information by cultural memory organizations for long-term public use. The Project grew, in part, from a panel discussion at the 2015 Depository Library Council meeting and a number of national meetings.

The PEGI Origin Story

At the October 2015 meeting of the Depository Library Council (DLC), a panel convened to discuss collaborative models of preserving federal government information. Panelists included librarians from California, Florida, Georgia, and South Carolina, as well as then-Dean of the University of North Texas Libraries, Dr. Martin Halbert. During this session, the discussion largely focused on preservation of physical collections and the conversion of tangible collections to digital format. Dr. Halbert, in contrast, spoke of the critical importance to advance an agenda of preserving electronic, i.e., born-digital, government information. As information shifts to a digital born environment, libraries and archives do not have the automatic fail-safes that they do for tangible materials, where items are received, made discoverable and then found and used by researchers. In most cases, librarians and researchers are not aware of materials produced and published solely on the web. Publishing directly to the web by federal agencies and commissions circumvents reporting and listing requirements within Title 44, Section 1902,1 thus making discoverability that much more difficult. During the DLC discussion Dr. Halbert noted, “It will be a challenge to cultivate a stronger consensus on the understanding of and importance of preserving digital government information.”

Dr. Halbert left the 2015 DLC meeting with the intention to build that stronger consensus and to engage stakeholders within libraries, archives, and research communities in conversation about the importance of preserving born-digital information and with the federal government about current policy and practice around information creation and possibilities for better preservation.

In conjunction with the 2016 spring Coalition for Networked Information (CNI) meeting, a group of librarians, archivists, technologists, and information professionals gathered for the Digital Preservation of Federal Information Summit (DPFIS). Lead by Dr. Halbert and the University of North Texas, the DPFIS used Collective Impact as a framework2 to engage national leaders in a structured, facilitated dialogue about at-risk digital government information and aimed to explore the development of a national agenda to address preservation and access of electronic government information. Outcomes of the summit revealed a shared concern among participants about the high risk of loss of electronic government information and a shared consensus about the need for an active coalition across public and private sectors to address preservation in this area. One surprising outcome of the summit was a lack of consistency in terminology. Despite participants being from similar and closely aligned professions, it quickly became apparent that each discipline defines some of the same terms differently.3 Terms like record, data, and publication differed greatly across professional practices.

Based on takeaways from the DPFIS, Dr. Halbert decided to convene another group of interested participants in conjunction with the 2016 fall CNI meeting. This much smaller meeting addressed some of the same issues, further elucidating the need for a shared vocabulary and a better understanding of current laws, practices, and preservation efforts. Additionally, this meeting brought together many of the individuals who would eventually join PEGI’s steering committee. At the closure of this meeting, participants reached consensus to pursue grant funding to conduct an environmental scan that would address the core set of themes and concerns common to the DPFIS meetings and to engage relevant stakeholders on broader needs and future implications of preserving electronic government information. It was decided to move forward as the PEGI Project.

Informing the PEGI Project

The PEGI Project is directed by its steering committee — Martin Halbert (now at the University of North Carolina at Greensboro); Roberta Sittel (University of North Texas); Marie Concannon (University of Missouri); James R. Jacobs (Stanford University); Lynda Kellam (University of North Carolina at Greensboro); Shari Laster (now at Arizona State University); and Scott Matheson (Yale University School of Law). The PEGI steering committee also includes Bernie Reilly and Marie Watzl, both of the Center for Research Libraries, and Deborah Caldwell, a graduate research assistant at the University of North Texas. The steering committee works in consult with staff at the U.S. Government Publishing Office (GPO) and the National Archives and Records Administration (NARA). We are primarily government information librarians, four of whom have served as chair of the Depository Library Council, and some have limited connections to digital preservation community. To this point, the PEGI Project is informed by the previously mentioned DPFIS meetings and, in part, by initiatives that have gained prominence in recent months, including DataRefuge and the Environmental Data Governance Initiative (EDGI).

Our committee consists of government information librarians who oversee Federal Depository Library Program (FDLP) collections, which are shifting in information transmission from print-centric to born-digital content. The committee is aware of outdated requirements and shortcomings of Title 44, which is the federal law that governs the GPO, the FDLP, Office of Management and Budget (OMB) and NARA. It is one of the goals of the Project to document and distill the policy and guidance around production, dissemination, and preservation of government information.

Members on the steering committee facilitate discussions that help to inform the Project on issues of infrastructure and workflows related to preserving electronic information as well as converting tangible content to digital formats. Additionally, all of the steering committee members bring experience and expertise in approaching projects and large initiatives collaboratively. The Project is identifying existing digital preservation partnerships and projects to understand efforts in this area and ways the PEGI Project can align with these efforts for greater collective impact.

The Project has engaged in a number of interviews with representatives across the federal government. PEGI Project team members met with employees from the Library of Congress Digital Preservation unit, NARA, the GPO, the Consumer Financial Protection Bureau (CFPB), the Marine Mammal Commission (MMC), and FEDLINK, the Federal Library and Information Network, which promotes consortia resources sharing among federal agencies. From these interviews, the team members learned more about the challenges of adapting print-centric law and policy to an electronic environment. Based on these interviews, the Project team is looking to identify additional agencies to consult regarding challenges and benefits with born-digital content as well as continuing conversations with NARA and GPO.

Activities and Outcomes of the PEGI Project

As mentioned above, the PEGI Project is a two-year initiative, with 2017 being the inaugural year, and year two concluding...
Ongoing throughout both years of the Project is an environmental scan that aims to identify aligned projects and existing digital repositories; gain a better understanding of policy and law around information production, dissemination and preservation; and describe differences in vocabulary and terminology across the library, archives and other aligned communities. Additionally, the Project team intends to continue to engage federal agencies and other identified stakeholders in one-on-one interviews. The interviews conducted during year one of the Project showed that federal agencies are interested in engaging more deliberately with the dissemination, access and preservation of their information products. The PEGI Project hopes to inform ways for all interested parties to convene more deliberately.

In conclusion, the PEGI Project aims to raise concerns and articulate solutions around the high risk for loss of electronic government information, and to address modes for preservation and permanent public access to born-digital government information. The library and archival community is, in some ways, ill-prepared for collecting, describing, and making available electronic government information. PEGI aims to identify current practices and potential efforts to create improved scenarios for future researcher communities to access tomorrow what is available today. To learn more about the PEGI Project’s objectives and activities, visit pegiproject.org.

Endnotes