PRESERVING AUTHENTIC ELECTRONIC RECORDS: THE INTERPARES PROJECT

BY ANNE GILLILAND-SWETLAND AND PHILIP EPPARD

In the introduction to his first book, *History of the Latin and Teutonic Nations* (1824), the German historian Leopold von Ranke made the famous statement that his history "seeks only to show what actually happened." Ranke launched a revolution in historiography by asserting the primacy of documentary research in archives to determine the facts about the past. His assumption was that using original archival records and manuscripts rather than secondary sources would increase the authoritativeness of historical writings.

Working with primary sources, however, requires that the scholarly community be able first to establish and corroborate the provenance, authority, and version of the texts with which they are working. While today's digital texts provide exciting opportunities to enhance scholarship because they can be made easily accessible online and can be searchable and easily manipulated, scholars and the archivists responsible for preserving these sources increasingly must contend with the intellectual ambiguity of records that are electronically created and maintained. It is difficult, using existing professional and scholarly methods, to establish the authorship, attribution, and versions of materials such as electronic drafts of documents or databases because the nature of the medium is so mutable. Furthermore, the inherent characteristics of increasingly prevalent electronic media forms make it no longer viable to consider preserving most such records in non-electronic form, since to do so would result in the loss of critical information as well as an equally critical loss of evidential value.

Ensuring the long-term preservation, trustworthiness, and accessibility of the nation's electronic records is a challenge that the National Historical Publications and Records Commission has been seeking to address for the past decade. The largest project that the Commission has supported in this endeavor to date is the InterPARES Project. InterPARES (International Research

InterPARES American research team members Ken Thibodeau, Sharon Farb, Anne Gilliland-Swetland, and Philip Eppard discuss matters at the project's Rome meeting. Photograph courtesy of Philip Eppard.

on Permanent Authentic Records in Electronic Systems) is a 3-year, multi-national research effort involving countries in North America, Europe, Asia, and Australia. Industry is represented in the research project through the participation of the Collaborative Electronic Notebook Systems Association (CENSA). The broad goal of the InterPARES Project is to develop the theoretical and methodological knowledge essential for the permanent preservation of records generated electronically and, on the basis of this knowledge, to formulate model policies, strategies, and standards capable of ensuring their preservation for use by archivists and other communities who need to retain and use trustworthy electronic records over indefinite periods of time.

The Commission has funded the non-Federal component of the American research team participating in InterPARES. The team includes researchers drawn from archival science, preservation management, library and information science, computer science, and electrical engineering, from the University at Albany, State University of New York; the University of California, Los Angeles; the Georgia Institute of Technology; the University of Missouri, Columbia; and Pennsylvania State University. The academic researchers are joined by representatives from the National Archives and Records Administration and the Smithsonian Institution. In addition to funding from the NHPRC, major funding contributions to the project as a whole have been made by the Social Sciences and Humanities Research

Council of Canada, the National Archives and Records Administration of the United States, and the Italian National Research Council. Other national archival institutions and universities in participating countries have also committed financial and research resources to the project. The international research team is under the general direction of Luciana Duranti, and the headquarters for the international project are located at the School of Library, Archives and Information Studies at the University of British Columbia.

While such models for conducting research are now commonplace in some other disciplines, InterPARES is the first example of a collaborative, multi-funded, multidisciplinary project emanating out of the archival community. The interest of so many governments, sectors, disciplinary domains, and archival institutions underscores both the ubiquity and the intractability of the challenges posed for the long-term management of valuable records that have been created and are maintained in electronic form.

The work of the project has been broken down into four research domains: I. Conceptual Requirements for Preserving Authentic Electronic Records; II. Appraisal Criteria and Methodology for Authentic Electronic Records; III. Methodologies for Preserving Authentic Electronic Records; and IV. Frameworks for Developing Policies, Strategies and Standards. Much of the research is being carried out by four task forces, corresponding to the four research domains. Several products have already been developed, and others are under development. These include requirements for establishing and preserving the authenticity of records in electronic systems; models of appraisal and preservation activities; a survey of institutions that are preserving electronic records or are involved in research in digital preservation; and a glossary of terms that are key to the findings of the research. The glossary will ultimately be

multilingual, and will also take account of variations in usage between different national and professional communities. While the glossary supports a full understanding of the products of the research, it is hoped that it will be useful to the broad archives, preservation, and digital library communities.

As part of Domain I., the project is in the process of conducting and analyzing extremely detailed case studies of electronic records systems of diverse types in a range of organizations. Many of these systems bear little resemblance to traditional forms of records (for example, complex databases, geographic information systems, laboratory records, and interactive web sites). From this analysis, we are deriving an understanding of the nature of the electronic record and of the extent to which its intellectual, if not its physical form, remains the same as that of traditional records. We are also building a typology of elements within different kinds of records that are crucial to the establishment and maintenance of the authenticity of that record while it is still current and when it becomes historical. This typology will then become the basis of the technical and policy requirements for preservation management systems and strategies.

Domains II and III are using modeling techniques to describe the components of the appraisal and preservation processes and also to analyze different methods and strategies currently in place or being developed in archival institutions. Domain IV is distilling the results of the work conducted in the other domains and addressing the policy and standards implications of these within different national and sector contexts.

The case studies have been conducted within government agencies of the United States, Canada, Italy, and the Netherlands; as well as at universities in the United States, Canada, and the United Kingdom; and in corporate settings in Italy. It is possible to compare similar types of recordkeeping systems that exist in different jurisdictions across the case studies, as, for example, in registry systems, student record systems, and Federal patent and trademark systems.

The work within Domains II and III will help introduce modeling methodology to the archival community. The task forces for these two domains have developed models that decompose values and functions of specific activities associated with archival appraisal and preservation. These models are currently being refined and reconciled with each other. What does a model do and how does it help archivists and systems designers? The fundamental role of modeling in the project is that it helps clarify the thinking of the research team. It ensures that concepts and activities and the relationships between them are defined precisely and consistently. The conceptual integrity and precision that modeling can give to the InterPARES research will also make it easier for systems designers to translate the models into working system. The preservation modeling is being informed by the results of an InterPARES survey of institutions that are preserving electronic records or are involved in research in digital preservation.

Members of the American research team have been actively disseminating information about the project and its work at meetings of the Society of American Archivists and meetings of several regional archival associations, but also at the conferences of other organizations, such as the American Society for Information Science and Technology. Many more presentations will be made and more products of the research will be released as the project moves into its final year. Further information about InterPARES is available at the project's web site, *www.inter-pares.org*, and at the web site for the American team, *www.gseis.is.ucla.edu/us-interpares*. The American team's web site also includes a comprehensive bibliography of the many publications, reports, and web sites of the other electronic records research projects funded by the National Historical Publications and Records Commission.

The dimensions of InterPARES are broad, but the challenges facing the long-term preservation of authentic electronic records are great. When future historians and other researchers seek to follow in Ranke's footsteps and "let the sources speak for themselves," the results of the InterPARES Project will help guarantee that they can do so with some level of confidence that the electronic records they are consulting are in fact the authentic and reliable sources required for their research.

ANNE GILLILAND-SWETLAND IS AN ASSISTANT PROFESSOR IN THE DEPARTMENT OF INFORMATION STUDIES AT THE UNIVERSITY OF CALIFORNIA AT LOS ANGELES. PHILIP EPPARD IS THE DEAN OF THE SCHOOL OF INFORMATION SCIENCE AND POLICY AT THE UNIVERSITY AT ALBANY, STATE UNIVERSITY OF NEW YORK.

Fynnette Eaton Joins Commission

The Society of American Archivists (SAA) has named Fynnette Eaton, Director of the Technical Services Division and Senior Electronic Records Specialist at the Smithsonian Institution Archives, to be its representative on the National Historical Publications and Records Commission. She succeeds Anne R. Kenney, Associate Director of the Department of Preservation, Cornell University Library, and Director of Programs, Council on Library and Information Resources (CLIR), in that capacity.

Ms. Eaton received a B.A. in history and an M.A. in English history from the University of Maryland. She began her archival career with the Archives of American Art, Smithsonian Institution, and subsequently took a position with the National Archives and Records Administration. Since 1986, when she joined NARA's Machine Readable Branch, Ms. Eaton has specialized in electronic records. As part of her work with NARA's Center for Electronic Records, she received the Interagency Committee on Information Resources Management's 1996 Technology Excellence Award for redesigning and implementing NARA's Archival Preservation System.

Ms. Eaton has been active in both the Mid-Atlantic Regional Archives Conference (MARAC), where she served on the Steering Committee and as Chair; and the Society of American Archivists, serving on several program committees, the editorial board for case studies in the management of electronic records, and the SAA Council. She was named a Fellow of SAA in 1995.