Presentation Title: Session introduction
Session: *The Long-Term Preservation of Authentic Electronic Records: InterPARES*
Event: Society of American Archivists Annual Meeting, Pittsburgh 1999
Presenter: Anne J. Gilliland-Swetland, Department of Information Studies, University of California, Los Angeles

The second half of the Twentieth Century has been marked in research as the era of *big science*. In the past two decades, international big science projects such as the Human Genome Project and the Upper Atmosphere Research Collaboratory have set the tone for the widespread development of mega-collaborative, multi-funded, multidisciplinary projects that are tackling issues that are often of critical concern to many national domains and disciplines, and that defy resolution by unilateral approaches. As archivists, we can look around today at allied disciplines and see many examples in non-scientific areas such as humanities text and digital library initiatives. The large-scale collaborative model has been promoted by funding agencies, governments, corporate institutions, and academic institutions, and adopted by researchers in both the public and the private sector for a number of reasons:

- the ability to bring diverse disciplinary perspectives to bear on resolving seemingly intransigent questions;
- the ability of technology to bring parties representing those diverse disciplinary perspectives together across time and space in a cost-effective way;
- the ability to have enough scope, granularity, and depth of expertise to parse large research questions into smaller constituent questions, and thus divide the research into more manageable, yet synchronized components;

and, equally importantly,

- the ability to generate the high profile for the research and its results that it often takes to get the attention of policy makers, standards developers, resource

allocators, academicians, and others who play key roles in translating the outcomes of the research into practice and ultimately into the fabric of daily life.

InterPARES has brought this model into the archival world by applying it to a set of issues relating to the preservation of authentic electronic records that are faced by governments, industry, and archival institutions worldwide, but that have failed to be satisfactorily addressed unilaterally by any of those sectors. If the organization of InterPARES seems complex as we explain it to our audiences, and the research goals and objectives seem multi-faceted and ambitious, it is because they are. In fact, some collaborative projects of this nature have so much complexity, and the issues surrounding multi-disciplinary collaboration are so substantive, that sociologists have been brought on board just to study the project's collaborative process and group dynamics. While we have not gone this far, it is certainly true that for the participants, the process of collaborating, and learning and honoring each member's national and disciplinary perspective is one of the great, invisible richnesses and perhaps less tangible outcomes of the project.

Our speakers today will be Terry Eastwood, co-investigator on the Canadian Research team; Maria Guercio, Director of the Italian Research Team; Hans Hoffman, Member of the Northern European Research Team; Rich Lysakowski, Director of the Industry Research Team; and myself, Co-Director of the American Research Team.

US-InterPARES

In a large-scale collaboration such as InterPARES, it is always critical that there be a unifying vision and set of research or development goals. At the same time, however, each participant must address the interests of his or her own constituents and funders. This is true for the national teams engaged in InterPARES, which are essentially functioning as distinct projects within the larger project, and are funded by different agencies with different priorities and sometimes even different timelines. It is also true *within* each national team, where members represent several different academic research interests as well as the interests of participating institutions such as, in the American case, the National Archives and Records Administration and various other testbeds in local government and college and university archives and records management programs. The US-InterPARES Project has capitalized on these inherent tensions by developing specific foci around several key issues and distributing responsibility for them according to team member expertise. The US-InterPARES members are drawn from academia and institutional archives across the country. They have expertise in a range of areas from archival science and preservation management to database design and mass storage media.

US-InterPARES is a three-year project that has been funded by the National Historical Publications and Records Commission (NHPRC). Although NHPRC funding only commenced in June 1999, American researchers have been involved with InterPARES since the planning stages. US-InterPARES researchers held their first working meeting in June in Bethesda, Maryland. For the duration of the project they will meet formally twice a year, alternating meetings between the east and west coasts, as well as less

formally at international team meetings and at SAA. A Website is available at http://is.gseis.ucla.edu/us-interpares which gives more information about US-InterPARES.

US-InterPARES, has three main sets of research and development activities which I would like to review briefly today. The first and overarching set of activities is to participate as the American Team in the international InterPARES Project, the scope and objectives of which have been already described by Terry Eastwood. The American team will work with a range of record types in legacy and active systems drawn from government and university settings across the United States. The systems on which the American Team has chosen to focus will be large database management systems (DBMS), which comprise the largest volume of electronic records in the Federal Government and thus present the most pressing electronic records questions for NARA; and multimedia electronic systems that look least like any analog equivalent and thus present some of the most tricky questions for the identification and preservation of authentic electronic records. Such systems would include geographic information systems (GIS), CAD/CAM, dynamic hypermedia (such as Web pages), and electronic data interchange (EDI) systems.

The second set of activities takes advantage of the computer science strengths of the US-InterPARES participants to conduct proof-of-concept work for InterPARES. These activities will occur in parallel with InterPARES developments and will involve translating draft theoretical models, templates, and typologies into systems design

requirements and data and metadata models. After analyzing the business process and resulting records systems, and designing and extracting relevant metadata, object-oriented or relational prototype systems will be built that incorporate data models and functional models. These prototypes will then be iteratively implemented and tested against relevant corpuses of records in different organizational, social, and national domains, and across different types of electronic records. By the end of 2.5 years, we expect that the prototypes for DBMS and GIS will be robust, and that we will also have developed and tested prototypes for additional types of recordkeeping systems, although these will not have gone through as many iterations.

Through this work, we will provide feedback that will then be used to reinforce or modify InterPARES' theoretical models, templates, and typologies, as appropriate. We also plan to generate and test systems analysis and design tools and strategies that can be factored into future systems design and policy and standards development in the areas of preservation of authentic electronic records in and over time.

The third set of activities falls under the broad rubric of translating research outcomes into practice and has several facets [overhead]:

- *Supporting and enhancing doctoral education and research in archival science:* Doctoral education and research is critical if we are to build the research base of the archival profession and to nurture the next generation of archival faculty. US-InterPARES was consciously designed to include significant support for doctoral

education and research, and explicit consideration of how such education and research might be enhanced. The students we are using have been handpicked because of their diverse disciplinary expertise and their research skills. The project, in turn, is providing them with extensive learning, research and teaching. The students are working closely with American and international researchers and their graduate students, and are also receiving training in diplomatics at UBC. It is anticipated that some of the students will conduct subsidiary research projects of their own whereby they will gather and analyze their dissertation data. Through their participation, doctoral students are getting an invaluable opportunity to see the international dimensions of archival theory and practice, and to apply a novel range of innovative research methods and analytical tools. In the course of US-InterPARES, we intend to assess the effectiveness of this research experience in preparing doctoral students for careers in archival research and teaching.

- *Master's education in archival science:* Archival faculty and doctoral students will develop curricular blueprints for graduate courses that integrate diplomatic concepts as well as the findings of this research into existing curricular frameworks and that could be taught individually or collaboratively between academic programs.
- *Continuing education for records creators, records managers, and archivists:* In addition to graduate education products, researchers will develop and pilot test continuing education and training products such as workshops, distance education courses, annotated bibliographies, glossaries, and potentially also case studies that

will then be made available for use by academic programs and professional associations and institutions in their continuing education programming;

- *Modeling and software tools:* A software resource group within US-InterPARES will develop modeling and object-oriented analysis tutorials and identify existing tools, especially those that are available at little or not cost. This group will develop computational models to test analytical and theoretical models, and possibly also modeling and analysis tools, all of which will be open source.