

## VI EUROPEAN CONFERENCE ON ARCHIVES

**31 maggio, mattina / May, 31 morning**

Sessione plenaria	Archivi elettronici: lo stato dell'arte nelle diverse realtà europee. Esperienze, scelte strategiche, normativa
Plenary session	Electronic Records: the state of the art in the various European countries. Experiences, strategic choices, legislation

**KENNETH THIBODEAU (NARA - National Archives and Records Administration, USA)**

***The Unsteady State of the Art of Preserving Electronic Records.***

The turn of the millennium has witnessed a significant shift in focus on issues related to the preservation of electronic records. In retrospect, one can characterize the decade of the 1990s as one which saw increasing concern with the growth in the use of computers for the production and communication of documents and with the challenges this growth created for managing, preserving and providing access to electronic records. Significant milestones reflecting this concern include the publication of three reports by the International Council on Archives' Committee on Electronic Records; the so-called "Pittsburgh Project" at the University of Pittsburgh in the United States; the collaboration between the U.S. Department of Defense and the University of British Columbia, Canada, to articulate formal requirements for the management of archival fonds and the subsequent establishment of the Department of Defense's standard for records management applications software; and the DLM-Forums sponsored by the European Union.

By and large, such efforts were analytic and dialectic. They explored the nature of the challenges posed by electronic records; contrasted the apparent implications of these challenges with traditional concepts, methods, and practices; attempted to determine requirements for addressing these problems; and proposed some methods for meeting such requirements. While there were, of course, some exceptions, the efforts of the 1990s were largely discursive: they discussed the problem but seldom got to the point of empirical work with electronic records either in a domain of research or practice. This discourse was undoubtedly essential, and in many respects it was both substantial and creative. However, it cannot be said to have been *productive* in the basic sense of that term: that is, in the sense of having produced practical implementations.

Clearly, solving the problems posed by electronic records requires moving from discourse to practice. This recognition is evident in many projects which started around the turn of the millennium. Many of these projects focus on what is the largest technological challenge posed by the spread of electronic records: that of preserving them across generations of rapidly changing technologies. This is the biggest technological challenge simply because the technology necessary to preserve most forms of electronic records being created either does not exist at all, has not been identified, has not been evaluated and tested empirically, or has not been shown to satisfy archival requirements.

One key contribution of the discourse of the 1990s derives from the fact that it did not drive towards practical solutions. This contribution is that it put us in a better position both to understand the problems we face and to identify and evaluate possible solutions.

A basic contribution is found in the ICA Guide to Managing Electronic Records from an Archival Perspective, which recognized that the preservation of electronic records is inextricable from their eventual communication; that is, we cannot be said to have preserved an electronic record unless we are able to communicate it in a form which is demonstrably authentic. Building on this recognition, the InterPARES project is exploring in depth what is required to deliver authentic electronic records over indefinite time through the work of its Preservation Task Force. Working on the foundation of the international reference model for an Open Archival Information System, the Preservation Task Force is developing a detailed formal model for such a system which addresses archival requirements for authenticity. In its analysis, the InterPARES Preservation Task Force has gone beyond the viewpoint expressed by the ICA Committee on Electronic Records to recognize that, not only is the communication of electronic records bound up with, and wholly dependent on, the methods used to retain them over time and technology, but that literally it is not possible to preserve an electronic record. It is only possible to maintain the ability to reproduce such a record.

This recognition has important and far reaching implications. The inevitable need to use hardware and apply software to reproduce a record from stored bits forces us to recognize that the problem of preserving/communicating authentic electronic records arises not over time, but at the first instance when a record is saved in digital form. There is always some possibility, when stored bits are retrieved and processed into a form wherein they can function as a record, that something may happen – intentionally or unintentionally – that will impact negatively on the integrity or authenticity of the record.

It is not that time is irrelevant. Continuing technological change over time makes the problem of delivering authentic electronic records increasingly difficult. Nonetheless, the problems archival institutions face in trying to preserve electronic records over the long term are not necessarily different than those faced by organizations managing active records over relatively short times. In fact, as governments and businesses prepare to shift more and more of their transactions to the Internet, they face *simultaneously* many of the same problems that archives face *chronologically* in trying to communicate electronic records. This reality forces us to supplement the traditional concept of an unbroken chain of custody of records with a parallel continuum: an unbroken chain of preservation. In the communication of electronic records across space, time, and technologies, any break in the chain of preservation may make it impossible to certify the authenticity of the records thereafter.

Such insight motivates the InterPARES Preservation Task Force to develop a model of the process of preserving electronic records which is much broader in scope than other approaches which focus on reactive steps that must be taken to counteract technological problems of obsolescence and media fragility. The InterPARES Preservation model seeks to identify every potential point of failure in the entire process of bringing electronic records into an archival system, maintaining them over time, and communicating them to users and to determine how such risks can be eliminated or mitigated. This approach does not prescribe techniques to counteract obsolescence or media fragility. Rather it provides a framework for evaluating such techniques. The criteria used for such evaluations derive from archival, institutional and societal requirements and constraints.

In a related effort, the U.S. National Archives and Records Administration has joined with the National Science Foundation to sponsor research efforts in the National Partnership for Advanced Computational Infrastructure to develop an information management architecture adequate and appropriate for managing, maintaining and

communicating authentic electronic records over time and technologies. This research seeks to derive robust and well-supported solutions to archival and records management problems from core technologies being developed in order to enable electronic commerce and electronic government. This initiative has the dual benefit of specifically and explicitly addressing archival and records management requirements and relying on technologies which have broad support, rather than on niche technologies developed especially for archives and records management. This initiative has the ambitious goals of developing an architecture which can be applied to virtually all types of electronic records, organized according to virtually any arbitrary structure imposed by their creators, including types of records and structures that have not been invented yet. It is intended to support all essential archival processes and to enable the application of special tools to handle unusual problems or specialized requirements. And the architecture is predicated on the principle that an archival information system must be essentially independent of the information technology which it uses at any time. That is, it must be possible to replace any component of hardware or software with minimum impact on other components and with negligible impact on the bodies of records being preserved. This infrastructure independence means that the archival system will not only be able to counteract obsolescence, but it will also be able to take advantage of improvements in technology over time. Research results to date have found that such an architecture must operate at three levels: data, information, and knowledge. Thus the methodology is described as "knowledge-based persistent object preservation."

Thus the state of the art of preserving electronic records is unsteady in two important respects: first, we are still grappling with the problem. We need to deepen our knowledge of its nature and complexity. We need to enrich our understanding both of what is needed and what is possible. Second, even when we have made substantial progress in these areas, we will still be faced with an evolving problem. We cannot seek a fixed solution to the problem of electronic records, because the problem itself will continue to change. To be viable over the long term, an archival solution must be inherently unstable or, to put it positively, it must be dynamic.