InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Case Study Proposal
Preservation and Authentication of Electronic Engineering and Manufacturing Records
Focus 2 - Science

Kenneth Thibodeau
National Archives and Records Administration

September 2003

Description

Various components of the U.S. Government are involved in the design, development, manufacture and use of complex systems. Many of these systems have very long life spans. In addition, there is an enduring requirement to retain records of the systems and their components. There may be need in the future to produce an identical system or to produce one or more parts for such systems. Such needs can only be addressed by retaining records which include both engineering specifications and manufacturing process methods. In the current environment, critical records, such as 3-dimensional models of solid artifacts and records of the process of manufacturing of such artifacts, exist and can only exist in digital form. Representing or reproducing a highly engineered system or component requires precise, accurate, reliable and authentic records. The Electronic Records Archives Program of the National Archives and Records Administration is engaged in a collaborative research project to develop and test prototype trusted computational environments for the creation, exchange and preservation of digital manufacturing models of machined piece-part shape and machining process knowledge. These models and process knowledge are records of manufacturing whose preservation, authentication and reliability are important for long term needs derived from the original purpose for which they are created.

In this research project, a trusted computational environment is defined as one having the capability to authenticate information entering or leaving the environment. Because of the diverse nature of the creation, exchange and storage environments, multiple perspectives on the requirements for authenticity will be encountered. The first objective of the project is to understand the conceptual, logical and physical requirements on the multiple abstract languages used to model part shape and manufacturing process knowledge. From this understanding, possibilities for message-based authentication will be explored.

This project involves partners in other U.S. Government organizations and in academia. Sample sets of electronic records will be generated by the records creator, transferred to the prototype archival system at the San Diego Supercomputer Center, subsequently retrieved by the creator, and then tested to determine their authenticity, reliability and usefulness.

The project is planned in two phases. The first phase will prototype the preservation, retrieval, and authentication of digital models of machined piece parts. The second phase will build on the results of the first, expanding the prototype to a significantly more complex model and to include records of the process which produced the system or part depicted in the model.

The prototype for both phases will involve all basic functions of an archival system as described in the OAIS standard, from submission by the producer to dissemination from the archives. The project will apply the findings of the InterPARES 1 project to support the authenticity of the records preserved in the prototype and will assess the reliability of the records to transmit required knowledge both of the systems and the processes over time.

Rationale

The records targeted in this case study fall within the InterPARES 2 intent to investigate "new record types and aggregations," with specific reference to the project's principal aim of "developing a theoretical understanding of the records generated by interactive, dynamic, and experiential systems, of their process of creation, and of their present and potential use in the artistic, scientific and government sectors." The records represent both new record types and new aggregations created in computer-assisted engineering, computer-assisted design and industrial automation systems. Little research has gone into the long-term preservation of authentic digital models of three-dimensional part shapes and the manufacturing process knowledge that is bound to them.

Of note in terms of the unprecedented nature of this case is the relationship between the records which model the system or part and the process records which are used in combination to produce the physical artifacts described in the model. Such systems are truly interactive, in a way that goes beyond the computer/user interactions described in the InterPARES 2 project proposal. In this instance, the electronic records are involved in the actual production of 3-dimensional artifacts.

The manufacturing prototype project explicitly builds on the results of the first InterPARES project. The U.S. Government has a long-term interest in the authenticity and reliability of the records derived the original purpose for which they are created. Thus the test case presents a prima facie instance of evaluating "methodologies for ensuring that the records created using these systems can be trusted as to their content (that is, are reliable and accurate) and as records (that is, are authentic) while used by the creator…." The research prototype project includes two institutions which are already

InterPARES participants: NARA and the San Diego Supercomputer Center (SDSC) and will test the persistent archives method developed by SDSC and NARA, applying advanced technologies in the test.

This case presents an unique opportunity in the government focus area to examine the reliability of electronic records: the ability of complex engineering records to stand for the solid objects modeled in the records, and of the manufacturing records to represent the processes required to produce such solid objects.

While the case focuses on a specific application within the U.S. Government, it has broad applicability worldwide. One aspect of this broad relevance is the role of standards in the research prototype. The U.S. case implements the ISO standard commonly referred to as the **S**tandard for **T**he **E**xchange of **P**roduct Model Data (STEP) and formally designated ISO 10303, Industrial Automation Systems and Integration – Product Data Representation and Exchange. NARA's government partners were, in fact, major contributors to the development of the STEP standard. The research prototype will also adhere to the ISO Open Archival Information System (OAIS) standard.

The relevance of the needs, issues and approach of the research prototype to other areas is reflected in independent international projects such as the Long Term Archiving and Retrieval of Product Data within the Aerospace Industry (LOTAR)[1] project in Europe and the Maturity Of Standard for Long-term CAD Data Archiving (MOSLA)[2] in Japan. Both of these projects start with the STEP and OAIS standards as defining a framework for addressing long-term primary needs for preservation of electronic records in the industrial sector.

The OAIS reference model was the basis of the InterPARES "Preserve Electronic Records" IDEF0 model. In the course of the development of the preservation model, the InterPARES Preservation Task Force interacted with the ISO working group responsible for the OAIS standard to promote consistency between the OAIS model, which addresses preservation of information in general, with the preservation model, which is specific to records. The case study of the research prototype will give InterPARES the opportunity to evaluate whether these two models are effectively synchronized.

Research Methodologies

The InterPARES case study team will directly observe the prototype archival system and the ingest, preservation and access processes executed in it., and will assess the case against the InterPARES 1 'Preserve Electronic Records' model. The research prototype itself will be modeled as part of the research project. The InterPARES team will formally assess the issues of authenticity, reliability and preservation of the records both with respect to InterPARES 1 findings and to address new issues and topics that emerge in the research prototype project.

[1] http://www.prostep.org/file/13270.WP_V10
[2] www.mosla.org

Reports of the research prototype project will be available to members of the case study team, and ultimately will be included in the documentation of the case study submitted to InterPARES. Members of the case study team will also be participants in the research prototype who will supply direct knowledge of the prototype. They will also solicit any additional information needed for the case study from other members of the research team.

Case Study Team

The case study team will be led Ken Thibodeau, Director, and include Fynnette Eaton and Michael Skipper of the ERA Program Management Office (PMO) at NARA. All are InterPARES researchers who will also be involved in the research prototype. In addition, in carrying out the case study they will interact with the principals working on the research prototype including Robert Chadduck, Director of Research, PMO, staff of NARA's federal partners and computer scientists and engineers at the San Diego Supercomputer Center.

Timeline

The research project's phase 1 will be completed by September 30, 2003, and phase 2 by September 30, 2004. The InterPARES case study of the project will start with the prototype project itself, but continue after the termination of the project to enable the InterPARES case study team to analyze the final results of the project. It is anticipated that the report of this analysis will be delivered to InterPARES by February, 2005.

Necessary resources

The resources required for the InterPARES case study of the manufacturing research prototype can be supplied by the government and academic organizations involved in the research prototype.