



InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Characterization of Case Study Validated

Case Study 19: Preservation and Authentication of Electronic Engineering and Manufacturing Records

Geneviève Sheppard, UBC
with input from Kenneth Hawkins

Version 1, October 2005

Relevance of the Case Study to InterPARES 2

The purpose of this case study is to examine “through an engineering experiment, the authentication of digital model (CAD) records using a content/message/semantic-based methodology rather than media, bit-count, or static provenancial attribute-based authentication.”¹ This study aims to examine the use of logic and semantics to query a digital entity’s meaning in order to assess its authenticity. The relevance of this case study is to enable InterPARES to reach its goals with respect to the study of scientific activities that are conducted using experiential, interactive and dynamic computer technology.

Information about the Creator

The Creator is defined as three separate research partners conducting the engineering experiment in which the creation, use, maintenance and disposition of the digital entities in the case study took place. These partners are: 1) an agency of the U.S. government with responsibilities in the science, engineering, design and manufacture of complex assemblies; 2) the Electronic Records Archives Program (ERA), National Archives and Records Administration (NARA); and, 3) The San Diego Supercomputer Center (SDSC). No information is provided in the case study final report with respect to the location and origins of the first research partner. The ERA is a program within

NARA, an agency governed by the National Archives and Records Act of 1984. The ERA is located in College Park, Maryland. The SDSC, at the University of California San Diego, was founded in 1985 through a grant from the National Science Foundation. The mandate of the first research partner is not specified in the case study final report. The mandate of the ERA is to “authentically preserve and provide access to any kind of electronic record, free from

¹ Case Study 19 Final Report, p. 1.

dependency on any specific hardware.”² The mandate of the SDSC is to “support international science and engineering discoveries through advances in computational science and high performance computing.”³

The function of the first research partner is to manufacture high-assurance, high tolerance machined piece parts for the U.S. government. The functions of the ERA include processing and storing electronic records and ensuring electronic records transferred to NARA remain free from corruption and are accessible regardless of changes in information technology. The main function of the SDSC is to operate powerful high-end computing resources to support research projects.

The structure of governance of the first research partner is not outlined in the case study final report. The ERA program falls within the governance of NARA and is managed by a Program Director and an Executive Officer. The actual system design of the ERA has been awarded as a contract to Lockheed Martin Corporation. The governance of the SDSC is through a Director, Executive Director and Division Director. The Center also has a separate director for each division, laboratory and department.

Information about the Administrative/Management Function

The first research partner is the initial creator of the digital entities serving as the objects of the study (though transformed at a later stage for the purposes of persistent archiving), and functions to produce CAD solid model files to be used in the design and manufacturing of mechanical piece-part assemblies. The creator must maintain these records authentic over time to “enable the production of the pieces as long as the business requires them.”⁴ The ERA program and SDSC Center participation in the experiment centers around archival experiment activities to explore persistent archiving of records in interactive, dynamic and experiential systems. The basis of the study comprises the abstraction of complex information from proprietary CAD formats, expression of this information into enhanced logical forms, rendering it into archival format, sending it across a trusted network to form part of a persistent archive and returning it for verification of authenticity, reliability and usability. The technical context and infrastructure of the experiment consists of the SDSC Storage Resource Broker (SRB) and metadata cataloguing system (MCAT) as well as the ERA Virtual Test Lab, which are all linked through a secure government network with limited access privileges.

Information about the Digital Entity Being studied

The digital entities being studied are the knowledge-enhanced objects derived from CAD files and STEP files for the purpose of persistent archiving. Product designers using proprietary Pro-Engineer CAD systems create the initial digital entities to aid in the design and manufacturing of mechanical piece-part assemblies. The proprietary CAD design records are then translated into Standard for the Exchange of Product Model Data (STEP) AP203 format. The logical form of the STEP records is enhanced into another logical form using C++ based knowledge representation tools. These entities are then taken through a proprietary reasoning engine (Logistica) into WC3 Ontologic Web Language (OWL) XML format.

The design engineer who creates the initial CAD model has no archive to persistently

² <http://www.archives.gov/era/about/welcome.html>.

³ <http://www.sdsc.edu/>.

⁴ Case Study 19 Final Report, p. 4.

store his work. An operational repository exists (Product Data Management System), but it does not offer the capability to provide access to the data, guarantee the authenticity of the records or ensure its usability in a computer application. When a solid model is registered in the Product Data Management System it is encapsulated with a STEP file as well a TIFF image of the drawing generated from the CAD model. The STEP file only contains the final model and does not provide information regarding construction technique. The creators of the models believe the construction file would be most useful to preserve, but it does not exist in any neutral standard form.