Scientific Data Curation in Government Agencies: Teaching Agency Data Creators How to Develop an OAIS-Compliant Digital Curation System

LORRAINE L. RICHARDS AND WILLIAM C. REGLI

Abstract: This poster will highlight a case study in which the Drexel College of Computing and Informatics (CCI) is collaborating with the Federal Aviation Administration (FAA) W.J. Hughes Technical Center (WJHTC) to develop digital curation capacity. PI William C. Regli and Co-PI Lorraine L. Richards, in conjunction with two doctoral-level research assistants, are helping the WJHTC develop capabilities to support enhanced data curation activities at the WJHTC by developing and enhance existing data sets and sources, mutually producing a "plan of action" for curating data from air flight scenarios, selecting an appropriate architecture, developing a prototype system based on that architecture, and creating requirements for a digital repository to serve the research needs of the Technical Center, Drexel University, and future users, including (eventually) the general public. The preservation functionality of the repository will be based upon the Open Archival Information System (OAIS) Reference Model. The proposed poster will address the steps, successes, and challenges associated with this project, which involves an organization that has a strong history of excellence in the creation and use of data for scientific purposes but which exhibits a relative lack of enterprise-level data curation and management processes. By attempting to move from a primarily manual and isolated, laboratory-based data management approach to an enterprise-level, trustworthy repository, the case of the FAA WJHTC illustrates the provision of a nontechnical curation/preservation education to data creators within a non-archival setting and the importance of building an appropriate curation infrastructure in a highly collaborative manner.

About the authors:

Dr. Lorraine Richards Bornn, Assistant Professor, performs research in the areas of digital curation and sustainable digital preservation. She is currently co-PI with Dr. William C. Regli on the Federal Aviation Administration-funded contract research project informing this poster – "A Research Study of Curation and Stewardship of Technical Data." She is also an instructor in the Digital Curation Professional Institute: Curation Practices for the Digital Object Lifecycle, hosted at the University of North Carolina at

Chapel Hill. She worked with the Blue Ribbon Task Force on the Sustainability of Digital Preservation and Access, coauthoring (under her former name "Eakin") their interim report. She has developed cost models for libraries and digital repositories, and developed the cost-model for the Dryad scientific data repository (Beagrie, Eakin-Richards, and Vision 2010). She also worked as an intern for the Council on Library Resources (CLIR), performing research activities for the Library of Congress. Richards Bornn was project manager of the IMLS-funded ESOPI-21 and ESOPI[2] projects led by Drs. Helen Tibbo and Christopher A. Lee at the University of North Carolina, Chapel Hill.

Dr. William C. Regli, Professor and Assistant Dean of Research, is a gregariously interdisciplinary researcher, with interests across computer science and engineering fields, including artificial intelligence, graphics, network-centric systems, and design and manufacturing. His research (Army, DISA, NSF, ONR, NRL, AFOSR, ARO, DARPA, NIST, NIJ/DoJ, FAA), has supported over 200 students and produced over 250 technical publications and 5 seminal patents in the areas of 3D/CAD search. He serves as Associate Dean for Research for Drexel's College of Computing & Informatics and Director of Applied Informatics – a 20-person RDT&E organization supporting public and private entities. Dr. Regli also serves as a Senior Science Adviser in the area of Engineering Informatics to the Defense Programs Office in the National Nuclear Security Administration of the US Department of Energy. Additionally, Regli co-directs the National Science Foundation's DataNet Federation Consortium which is creating a trans-disciplinary, national-scale infrastructure for science and engineering data stewardship. Dr. Regli will be serving on a two-year "loan" to the Defense Advanced Research Projects Agency (DARPA) in the role of the Deputy Director of the Defense Sciences Office (DSO), beginning September 2, 2014.