

add-ons for OpenOffice/StarOffice, FrameMaker and MS Office. They especially exploit the XML support most software vendors of office systems have newly integrated as a genuine standard interface into their products. The generated XML files form the source for different presentation formats, the basis for longterm preservation activities, and a prerequisite for value added retrieval techniques.

The paper will reveal both, the DTD generation system and the XML conversion process.

**Title: Avoiding the Digital Dark Age with PDF & XML: How your institution can ensure access to digital theses and dissertations and avoid technological obsolescence**

**Authors: Chuck Myers (Adobe Systems)**

**Abstract:** If you suddenly couldn't access the building where your valuable research data or study was kept, what would you do? In a paper-based world, archiving graduate research meant storage of paper or microfilm. But what happens as more and more theses and dissertations are created electronically? How do you preserve valuable research in a consistent format? How do you keep the exact look and feel of a document today, 30 or 300 years from today?

The lack of a recognized and accepted electronic standard for archiving theses and dissertations -- particularly as new generations of hardware and software make previous digital technology obsolete - could lead to the loss of significant amounts of valuable information. For example, over the past several decades, military files from the Vietnam War, records from the Viking Mars Mission, Census Bureau data and land use records have been lost due to the inability to read data formats and the deterioration of magnetic tapes used to store that data.

Archiving standards such as PDF/A and PDF/XML can ensure secure access to digital data across the enterprise and Internet. The speaker will discuss how graduate schools could use PDF/A and PDF/XML to archive and preserve their digital works enabling the fidelity of digital documents to be preserved for generations to come. He will examine how XML as a neutral format for creating Web, print, and wireless content can be formatted and presented as a PDF document. The session describes the new place of XML and Web standards for printable representation of dynamic text and graphics as well as static content. He will discuss how these standards are being developed under the publicly available and highly prevalent PDF standard, offering protection from technological obsolescence over the ages.

Attendees will learn about:

- The dramatic growth of information due to the digital age and the challenges involved in preserving digital data.

- Examples of how universities are using PDF and could use PDF/A for preserving digital theses and dissertations.
- How PDF supports XML for ensuring reliable presentation of documents
- Different XML languages and formatting tools
- How to create personalized documents using XML and SVG (scalable vector graphics) for presentation within the PDF format
- The security features and universal interoperability that have earned ISO endorsement of PDF as an open standard for electronic documentation.
- How to use PDF to convert Web pages and other electronic content to digital versions that have a small footprint yet retain the complete look of the original.
- How to use PDF as an “electronic envelope” for storing other file formats within a PDF container.
- How PDF/XML enables institutions to embed XML data in PDF for reliable archiving, sharing, viewing, and interacting through Acrobat 6.0 and the Adobe Reader

**Title: Theses and Orations in the Digital Academic REpository (DARE) of Universiteit Leiden**

**Authors: Trudi C. Noordermeer (Universiteit Leiden)**

**Abstract:** Universiteit Leiden was founded in 1572. The university trades PhD theses with other European universities since the end of the sixteenth century. In The Netherlands several universities publish PhD theses and orations on the Internet since 1997. In 2004 Universiteit Leiden carries out a pilot project with the objective to publish PhD theses and orations on line and to investigate the feasibility of a long term operational service. The project concerns the entire workflow:

- production by the authors;
- communication with the authors;
- selection criteria and acquisition procedures;
- authenticity;
- metadata according to OAI-MPH;
- storage in the Digital Academic REpository of Universiteit Leiden using Digitool from Ex Libris or DSpace;
- on line publication on the Internet, if permission is received;
- copyright issues;
- interoperability with the infrastructure of other Dutch universities which cooperate in the national DARE project, organized by Surfnet, the national computernetwork for higher education and research in The Netherlands;
- feasibility of re-use of parts of the theses in the e-learning environment Blackboard;