Towards support for long-term digital preservation in product life cycle management

Wolfgang Wilkes
Jörg Brunsmann
Dominic Heutelbeck
Matthias Hemmje

Andreas Hundsdörfer
Hans-Ulrich Heidbrink

FernUniversität in Hagen
InConTec
Shaman Project

• Shaman: Sustaining Heritage Access through Multivalent Archiving

• EU integrated project (IP)
  • 15 partners from 7 European countries
  • Duration 2008 - 2011

• Goals
  • Development of a long-term next generation digital preservation framework
  • Application in 3 domains
    – Libraries and archives
    – E-Science
    – Engineering environments
Outline

• PLM and Engineering, why DP in this domain?
• Specific problems of DP of PLM data
• Shaman approach
  – Definition of a preservation information life cycle
  – Architecture for interaction between PLM and DP
• Conclusion and outlook
Domains / Characteristics

• Various industries and disciplines involved
  • Automotive
  • Aerospace
  • Process plants
  • Equipment suppliers
  • Mechanical engineering
  • Electronic engineering
  • Software engineering
  • …
Why digital preservation in this domain?

• Long living products, e.g.
  • Aeroplanes
  • Cars
  • Process plants

• Legal requirements for archival
  • Laws
  • Contractual requirements
  • Preparation for legal defense

• Economic requirements for digital preservation
  • Product improvements and adaptations
  • Reuse of product designs
Specific problems of DP in Engineering

• Complex, structured data
  • Various kinds of data
    – Requirements, logic, layout, geometry, manufacturing data, etc.
  • Tool based data generation
    – Often proprietary data formats
  • References to external sources

• Life cycle phases and development processes
  • Ingest becomes a process – not a single operation
  • Processes and related data provenance has to be captured
  • Cross company collaboration → IP disclosure concerns
Shaman approach to DP and PLM

- Focus: Interaction of PLM system and DP system
- Definition of a DP information life cycle
  - pre-ingest
  - post-access
- Architecture for interaction between PLM and DP
  - Integration of DP activities in PLM processes
Shaman information life cycle

- Pre-ingest
  - Creation
    - capturing of additional meta data
    - transformation to standards + validation
  - Assembly
    - Collection of data to be preserved

- Post access
  - Adoption
    - Transformation of standards to tool specific formats
    - Validation of data correctness
  - Reuse
    - Integration of preserved data into actual designs
Summary and conclusions

• Both PLM and DP systems are extended by specific functionality
• DP service interface
• PLM additional functions for
  • Capture of additional meta data
  • Selection of data to be preserved
  • Transformation from and to standard representations
  • Validation of accessed data
• DP additional functions for
  • Access control, license management
  • Validation of ingested data
  • Semantic, meta data based search functionality
Some open research topics

• Detailed specification of the DP service interface
• Distributed archives
• Capturing and generation of meta data
• Dependencies to external ontologies and knowledge sources with independent evolution processes
Any Questions?

Wolfgang Wilkes
wolfgang.wilkes@fernuni-hagen.de

Jörg Brunsmann
Dominic Heutelbeck
Matthias Hemmje

Andreas Hundsdörfer
Hans-Ulrich Heidbrink