

# eInfrastructures & DL

**... the future**

**Matthias Hemmje**, University of Hagen, Germany

**TPDL CONFERENCE**

**Malta 24<sup>th</sup> of September 2013**

# Overview

- SHAMAN
- Parse.Insight
- APARSEN
- SCIDIP-ES

Securing  
Communication with  
the future



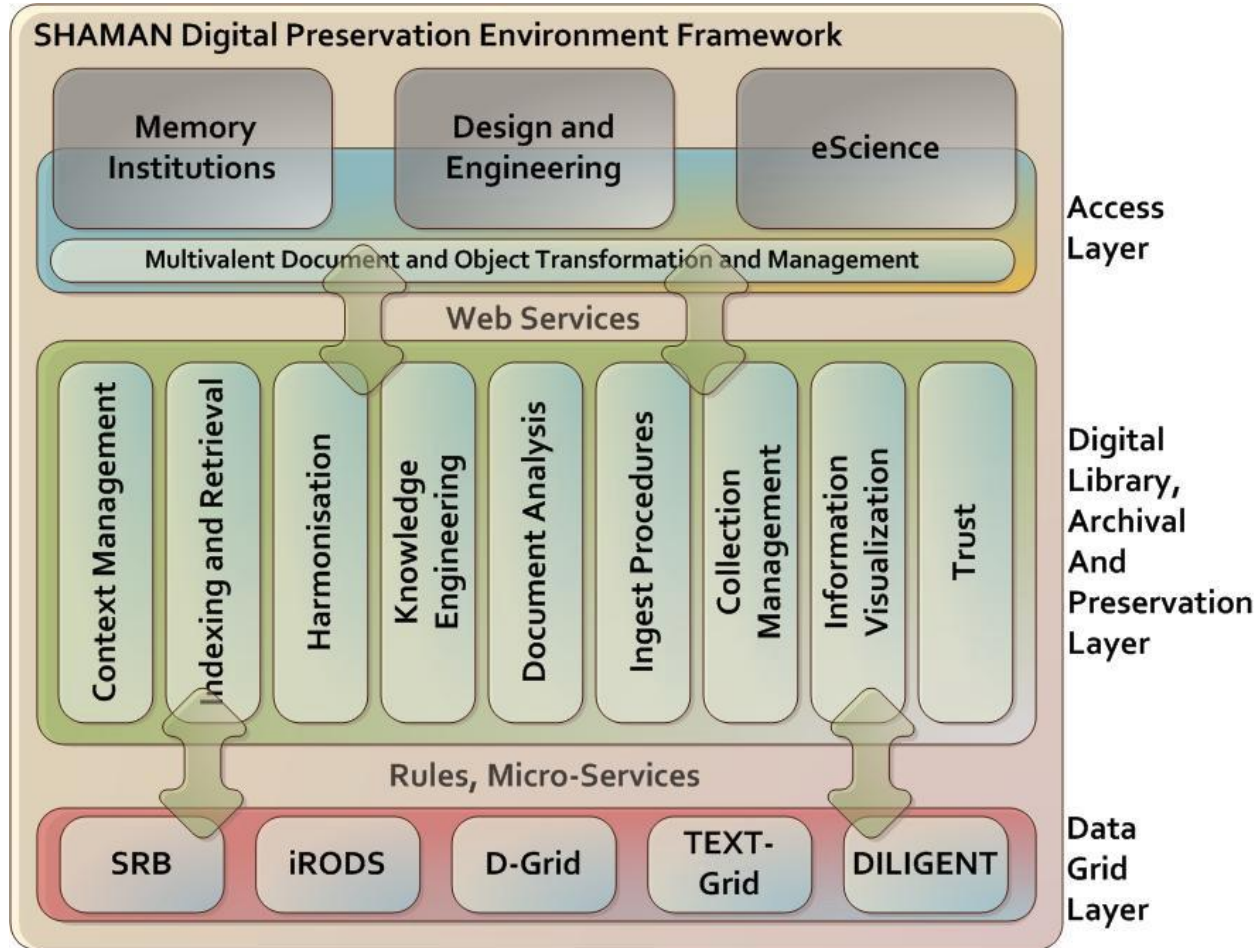
# SHAMAN

Sustaining Heritage Access through Multivalent ArchiviNg



Co-funded by the European Union

# SHAMAN R&D (I): Conceptual Architecture & Grid Utilization



# SHAMAN R&D (II): Supporting Storage and Curatorship – Distribution vs. Centralization

## Organizational aspects

- the SHAMAN framework can support centralized as well as distributed collection **storage**, indexing, and analysis as well as hybrid forms of these
- the SHAMAN framework can support centralized as well as distributed **curatorship** as well as hybrid forms of both
- All the above mentioned combinations require appropriate global as well as local policies and mutual agreements on the organization level

## Possible implications

- the SHAMAN framework can equally well drive application solutions which are tailored towards user community requirements of
  - storage and access distribution as well as
  - individual organizational & curation policies as well as
  - relationships amongst collaboratively curating institutions,
- this means the SHAMAN framework can help bridge gaps between individuals, teams, institutions/organizations, and working cultures within DP application domains

# SHAMAN R&D (III): Enabling Migration, Emulation & Hybrid Approaches

**SHAMAN Supports Migration through defining preservation policies which describe:**

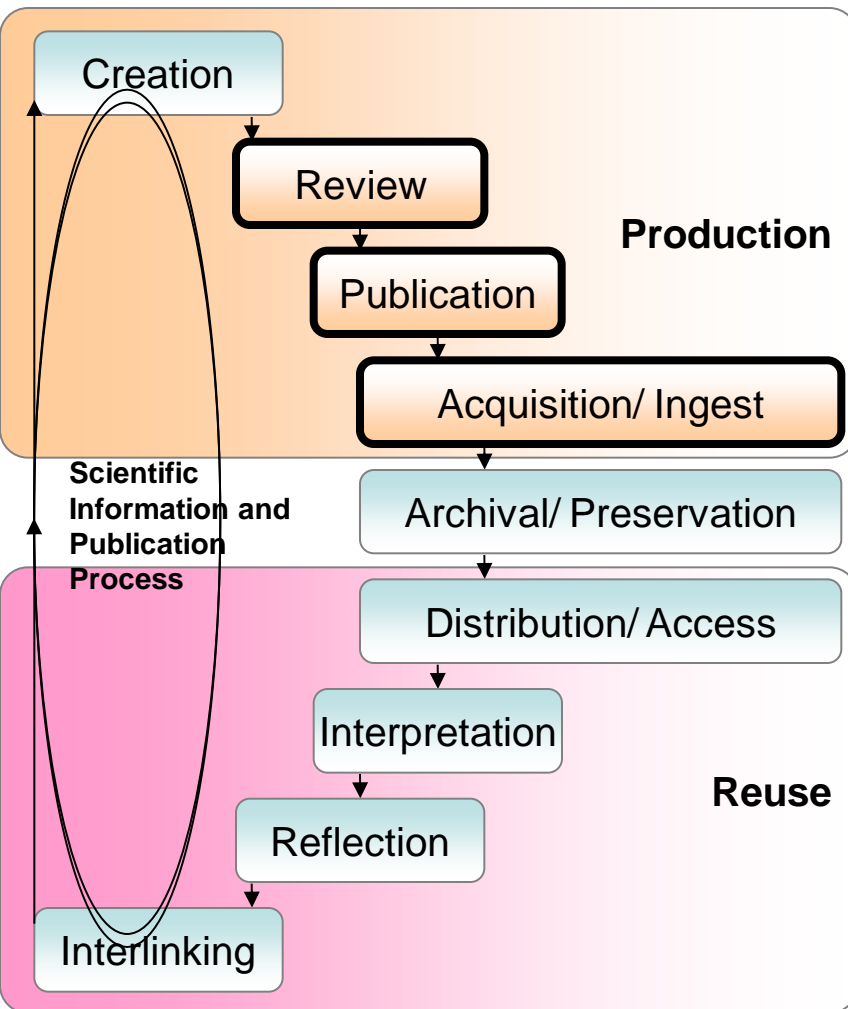
- the objects of the sub-collection to be migrated
- rules by which the object formats are migrated
- the relevant set of formats related to the meta-data and process contexts
- rules by which the meta-data and process formats are migrated

**SHAMAN Supports Emulation policies through defining preservation policies which describe:**

- the display technologies of the objects of the sub-collection to be migrated
- the access&re-use technologies for the relevant meta-data contexts
- the access technologies for the relevant process contexts
- rules by which these technologies are emulated

**In addition, SHAMAN Supports Hybrid Approaches, too.**

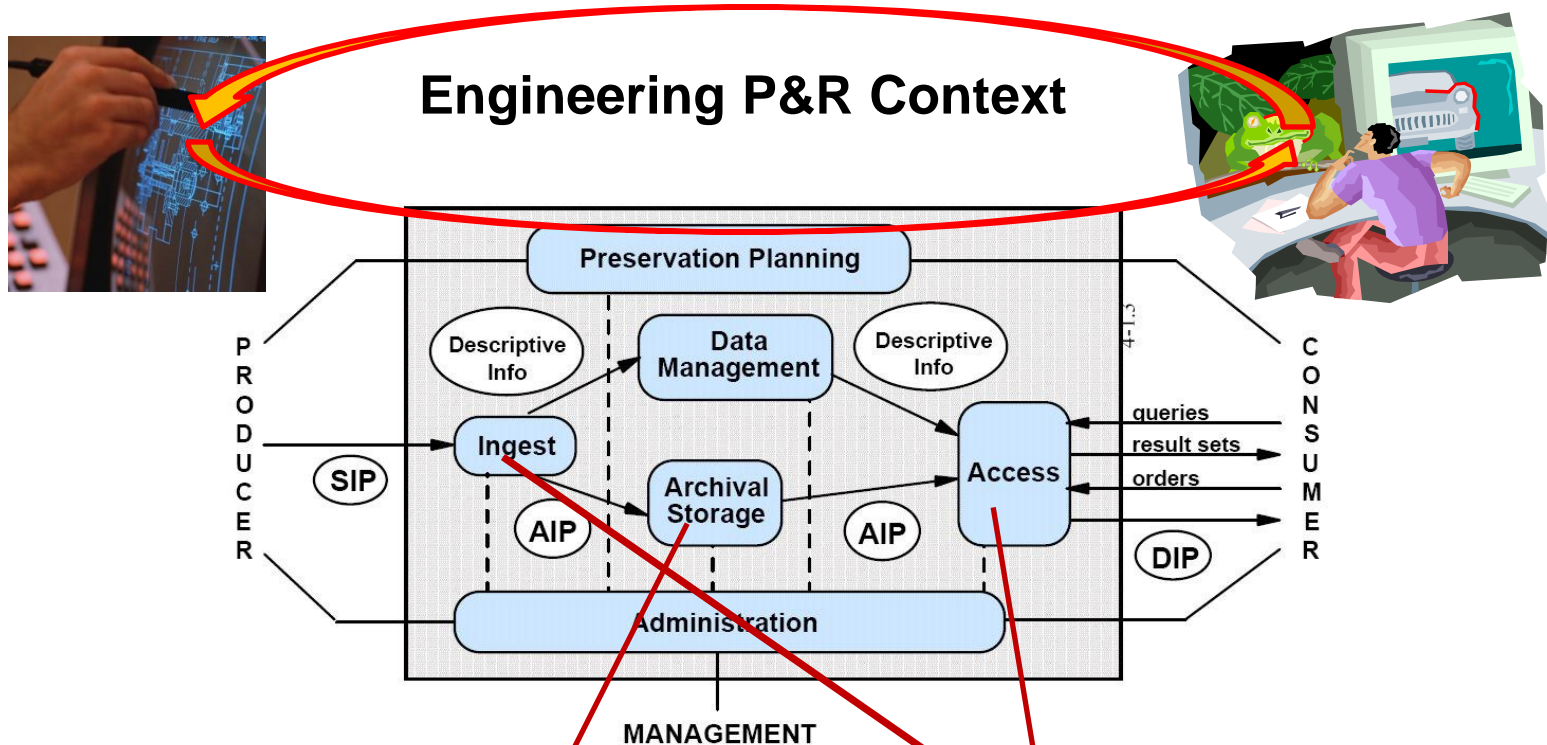
# SHAMAN Context Model: Use Case Scientific Publication



- Scientific Publication process can make available rich set of information to the *reuse context*
- Scientific community web publishing, information systems, repositories and further DL applications can be extended to capture context data beyond the immediate requirements of scientific publishing



# Additional SHAMAN Use Case Examples: Engineering Scenario and the OAIS Reference Model



Multivalent approaches preserving CAD models

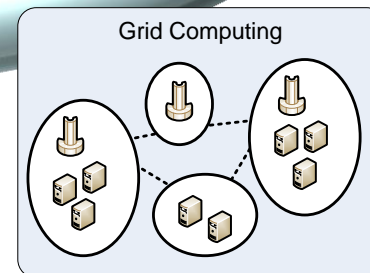
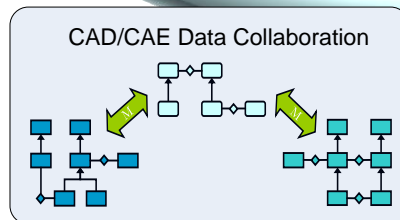
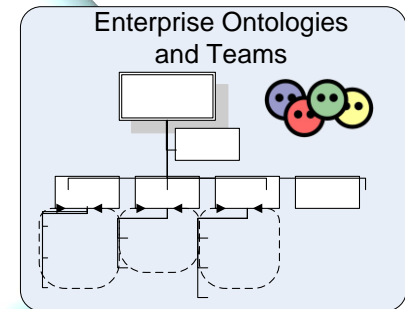
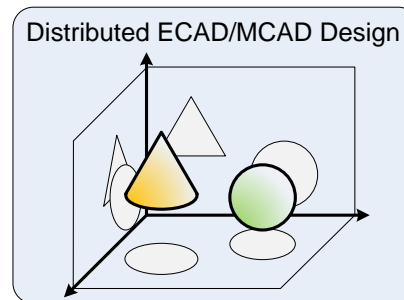
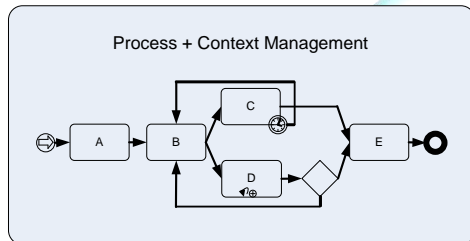
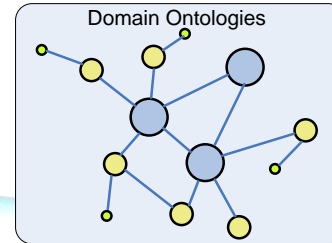
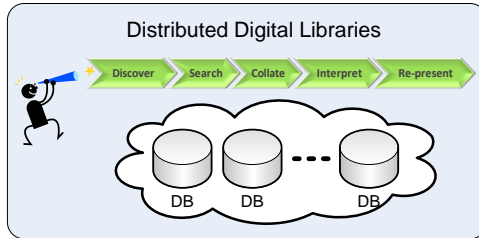
- migration
- emulation

Providing ingestion and access in distributed heterogeneous industrial engineering szenarios with collaboration support

Base diagram from: Consultative Comitee for Space Data Systems (2002):  
*Reference Model for an Open Archival Information System (OAIS)*; CCSDS 650.0-B-1; BLUE BOOK



# Additional SHAMAN Use Case Examples (II): R&D Dimensions of the SHAMAN Engineering Scenario





# Science Data Preservation Infrastructure Roadmap

# Achievements

- Roadmap development
  - Review and synthesize existing roadmaps
  - Produce draft on this basis
  - Refine in the light of evidence
  - Used to provide some structure to Surveys
    - BUT needed to be careful to avoid pre-determining the responses

# the Parse.Insight Roadmap

- Alliance Research programme
- Alliance action plan
- CASPAR Conceptual Model
- JISC Digital Repositories Roadmap
- DCC Lifecycle model
- DPE Research Roadmap
- Warwick workshop report: Digital Curation and Preservation: Defining the research agenda for the next decade
- Requirements for Digital Preservation Systems: A Bottom-Up Approach (2005)
- Thirteen Ways of Looking at...Digital Preservation (2004)
- Overview of Technological Approaches to Digital Preservation and Challenges in Coming Years (2002)
- Digital Preservation and Deep Infrastructure (2002)
- Report of the Task Force on Archiving of Digital
- Mind the Gap report from the UK DPC
- E-INFRASTRUCTURE STRATEGY FOR RESEARCH: FINAL REPORT FROM THE OSI PRESERVATION AND CURATION WORKING GROUP
- eIRG Roadmap
- ESFRI Roadmap
- Developing World-class Research Infrastructures for the European Research Area (ERA) - report of the expert group
- CASPAR test case questionnaire
- Cyberinfrastructure Vision for 21st Century Discovery
- Invest to Save
- It's About Time
- Stewardship of Digital Research Data: A Framework of Principles and Guidelines
- To Stand the Test of Time - Long-term Stewardship of Digital Data Sets in Science and Engineering

# Significant Results

- Components of roadmap
  - Financial
  - Organisational/social
  - Policy
  - Technical
  - (plus virtualisation as underlying concept)
- Importance of OAIS for structuring of roadmap

# Use of the Roadmap

- Guide networking activities
  - APARSEN Network of Excellence bid
- Basis of technical design of preservation infrastructure in SCIDIP-ES
- Useful support for High Level Expert Group report
- Supporting evidence



Tools

COMMON VISION

Digital Preservation activities

Information about users and practices



Infrastructure



scidip-es

ISO : Audit and Certification Standards plus organisation

ISO standard: OAIS update

ISO standard: OAIS

Relationship to related work and community practices



# Fine

Thank you very much for your attention.