Digital Preservation of Business Processes and Services
April 2011 - December 2014

Angela Dappert
DPC
angela@dpconline.org
Overview

- Project
- Goals
- Approach and methodology
Digital preservation for timeless business processes and services

- timbusproject.net/
- http://opensourceprojects.eu
- info@timbusproject.net
- https://twitter.com/timbus_project

- April 2011 - December 2014

- co-funded by the European Union FP7/2007-2013 under grant agreement no. 269940
The TIMBUS Consortium

TIMELESS BUSINESS

- SAP - Lead partner (NI, CH)
- Intel (Ireland)
- Software Quality Systems (Germany)

- Digital Preservation Coalition (UK)
- INESC - ID (Portugal)
- Karlsruhe Institute for Technology (Germany)
- Laboratório Nacional de Engenharia Civil (Portugal)
- Münster University (Germany)

- Caixa Magica Software (Portugal)
- Secure Business Austria (Austria)

Industry
Research
SMEs
Overview

- Project
- Goals
- Approach and methodology
A Preservation Continuum

Digital Preservation

- Files
- Data
- Hardware
- Software
- Processes
- Services
- Semantics
- Representation Information

Complexity
Business Continuity Management

Longevity

3rd party
Motivation: Continuity

- Reusability of any part of the process:
  - Redeployment of the process
- Curation and continuous improvement
  - By explicitly exposing the methods used in the process
- Changes in technical environments
  - Manage transitions across platforms
- Staff changes
  - Knowledge retention and continuity
- Third-party services
  - Protect data and functionality
Motivation: Provenance

- Improve understandability
  Explicit links between data, processes, platforms, results

- Attribution and referencing
  Cite data, publications and processes using URIs/DOIs
  Refer to registries

- Prove of authenticity

- Repeatability /Reproducibility

- Traceability, error detection, diagnosis and prove of quality of process products
  Assess validity of data, processes, derived outputs

- Regulated areas
  Audit, compliance, reproduce, diagnose, evidence for litigation
Use Cases

Preservation of

- Open source workflows
- Open source systems
- Civil engineering: Dams
- E-Health
Overview

- Project
- Goals
- Approach and methodology
Context Description

Customer Environment

Software
Data
Hardware
Services

Semantics
Representation Information

3rd party

Context Model

OWL
DIO
OWL
DSOs
Context Acquisition

Customer Environment

Context Acquisition
- Processes
- Properties
- Dependencies
- ...

Process Extractors
Process Monitoring
Modified Code
Network Information
Network Topology
Network Monitoring

Context Model
Instance

Edit Context Instance

Data Representation Information
- Files
- Data
- Software
- Hardware
- Services
- Processes
- 3rd party

Semantics

- Windows MSI
- Perl Modules
- (LLM Surveys)
Context Analysis

TIMELESS BUSINESS

Domain Experts:

Preservability

Legalities
Life-Cycle
Management

Risk Expert
Console

Reports:

- Problems
- Opportunities

Risks:

- Identify
- Classify

Risk Evaluation
& Treatment

OWL Context
Model Instance

22 September 2014

timbusproject.net © 2011-2014
## Context Analysis - Preservability


<table>
<thead>
<tr>
<th>Quality</th>
<th>No.</th>
<th>Characteristics</th>
<th>Metrics</th>
<th>Expected Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compatibility</strong></td>
<td>C1</td>
<td>Co-Existence</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>C2</td>
<td>Interoperability</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Portability</strong></td>
<td>P1</td>
<td>Adaptability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P2</td>
<td>Installability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>P3</td>
<td>Replaceability</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maintainability</strong></td>
<td>M1</td>
<td>Modularity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M2</td>
<td>Reusability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M3</td>
<td>Analysability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M4</td>
<td>Modifiability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M5</td>
<td>Testability</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Security</strong></td>
<td>S1</td>
<td>Confidentiality</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Context Analysis - Preservability

### Quality

<table>
<thead>
<tr>
<th>Quality</th>
<th>No.</th>
<th>Characteristics</th>
<th>Metrics</th>
<th>Expected Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Compatibility</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>Co-Existence</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>Interoperability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Portability</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P1</td>
<td>Adaptability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P2</td>
<td>Installability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P3</td>
<td>Replaceability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Maintainability</strong></td>
<td>M1</td>
<td>Modularity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M1.1</td>
<td></td>
<td>The system's coupling is low and cohesion is high.</td>
<td>Coupling and cohesion metric values</td>
<td></td>
</tr>
<tr>
<td>M2</td>
<td>Reusability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M2.1</td>
<td></td>
<td>The external interfaces of the system are clearly specified.</td>
<td>See C2.4.</td>
<td></td>
</tr>
<tr>
<td>M2.2</td>
<td></td>
<td>The communication to and from the system is standardised.</td>
<td>See C2.2.</td>
<td></td>
</tr>
<tr>
<td>M2.3</td>
<td></td>
<td>The system has a sufficient level of encapsulation.</td>
<td>See M1.</td>
<td></td>
</tr>
<tr>
<td>M2.4</td>
<td></td>
<td>The licenses and legal regulations permit reuse.</td>
<td>License clauses</td>
<td></td>
</tr>
<tr>
<td><strong>Security</strong></td>
<td>S1</td>
<td>Confidentiality</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Context Analysis - Preservability

- **loose coupling tight cohesion**

22 September 2014 timbusproject.net © 2011-2014
Context Analysis

Domain Experts:
- Preservability
- Legalities Life-Cycle Management

Reports:
- Problems
- Opportunities

Risks:
- Identify
- Classify

Risk Expert Console

Risk Evaluation & Treatment

OWL Context Model Instance

Risk Log
Context Analysis

Domain Experts:
- Preservability
- Legalities Life-Cycle Management

Risk Evaluation & Treatment

Risks:
- Identify
- Classify
- Mitigate

Solutions:
- Escrow Service
- Digital Preservation Actions

OWL Context Model Instance

Risk Expert Console

Domain Experts' Actions:
- Identify
- Classify
- Mitigate

Solutions' Actions:
- Risk Recommendation + Cost

Digital Preservation Actions' Actions:
- Risk Recommendation + Cost

Preservability Context Report
Legalities Life-Cycle Management Context Report

Risk Expert Console

Technical Expert Console

22 September 2014

timbusproject.net © 2011-2014

18
Putting it Together

TIMELESS BUSINESS

Legalities
Life-Cycle Management

Preservability Qualities
Domain Experts

Customer Environment

Context Acquisition
• Processes
• Properties
• Dependencies
• ...

Context Model Instance

Risk Evaluation & Treatment

Risk ➔
Cost + Recommendation

Virtualise, Validate, Store
Monitor
Redeploy

Escrow Services

Digital Preservation Actions
Solutions

Domain Experts

Risk ➔

Risk Evaluation & Treatment

Context Model

Context Model Instance

Data
Representation Information

Softwar e
File s

Hardware

Processes

3rd party

Semantic

Monitor
Redeploy
• Determine artefacts to be preserved
  ▪ Calculate cost of preservation based on all the required artefacts
  ▪ Calculate projected cost for a given number of years
  ▪ Recommend action

• Package artefacts
  ▪ Create list of artefacts to acquire from the context model
  ▪ Create checksums for all files
  ▪ Extract system metadata for the files to store with the files
  ▪ Create metadata for the overall project to store with the project
  ▪ Annotate the context model with the newly created metadata
  ▪ Specify significant properties, their known values, how to capture them

• Create a Virtual Machine
  ▪ ... using the hardware configuration extracted during context acquisition
  ▪ Verify the service
  ▪ Store VM in backend storage
  ▪ Update risk register with residual risks