Understanding Metadata

A Process Oriented Approach to Recordkeeping Metadata

Peter Horsman
Archiefschool, Amsterdam
Confusion

- Concept of Metadata
- No embedding in recordkeeping and archive theories
- Orientation on daily practice
- Not used to Models
- Complexity of Metadata schemas

- Aversion
Contents

- Hypothesis
- Metadata categories
- Methodologies for systems design
- Method for metadata analysis
- Techniques: user understandable modelling
- Validation
- Conclusion
Hypothesis

- A metadata system is not a goal in itself
- Implementation of a theoretical metadata schema is not a practical approach
- Metadata are closely linked to recordkeeping processes
- Processes are better understood than metadata
Metadata categories

- Descriptive – Intellectual Control
  - contents, context, structure
- Administrative – Administrative Control
  - provenance, ownership, rights, access, preservation
- Structural – Physical Control
  - technical
Metadata creation

- Deliberately created / captured by recordkeepers
  - Retrieval
  - Interpretation
- ‘By products’ of recordkeeping activities
  - Created and used in recordkeeping processes
  - Accountability
# Metadata Matrix

<table>
<thead>
<tr>
<th></th>
<th>Intellectual control</th>
<th>Administrative control</th>
<th>Physical control</th>
</tr>
</thead>
<tbody>
<tr>
<td>By-product</td>
<td>X</td>
<td>XX</td>
<td>X</td>
</tr>
<tr>
<td>Description</td>
<td>XX</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Capture</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Systems Design

Assumption:

- A Recordkeeping system is an Information system
- Development of RKS is IS development
- Methodologies for IS development applicable for RKS development
System Design

- Defining the System
  - RKS As a Black Box
  - Functions, Purpose, Interactions with Environment

- Analysing Processes
  - Identification, analysis, inputs/outputs, dataflows
  - Controls

- Analysing Data

- Linking Processes with Data

- User Involvement
  - Processes better understood than data
System Design

Process Analysis

- Identifying Recordkeeping Processes
- Identifying Controls
- Identifying Stakeholders and experts
- Identifying existing documentation
- Interviews
- Business Model as a framework / starting point
- Process decomposition – linking processes through flows (inputs and outputs)
RKS business model

Management and Control

Ingest

Preservation

Records

Metadata system

Processing

Access

Search, Retrieval and Use
System Design

Process Analysis: Example Ingest

- Input: Records / documents – From?
- Input: Metadata (in different terminology) –
  From business application, other system
- Output: Captured records – To Storage / Process
- Output: Accession registration (local standards) / structural metadata – To metadata store
- Controls: Criteria, Forms, Rules ...
System Design

Documentation and Feed-back

controls

INGEST PROCESS

Records

Metadata

Metadata Store
Metadata Analysis

- Through analysis of processes
- Identifying Entities and data elements
- Terminology Trap
- Metadata Model
- Cross reference with Processes
- Documentation
Validation

- Recordkeeping Metadata schema as a Reference Model
- Identifying gaps
- Expert feedback
- Differences in terminology
Conclusion

- Records professionals primary users of metadata systems - experts
- User involvement indispensable
- Metadata must be understandable for users
- Metadata schemas useful tools for analysts
- Labour intensive, but: process driven metadata analysis pays back