Maarten Hoogerwerf & Mike Priddy
Data Archiving and Networked Services (DANS)
IASSIST 2013, May 30th, Cologne.

A Reference Architecture for Research Infrastructures in the Social Sciences & Humanities
Overview

• Research Infrastructures - stakeholder perspectives
• What is a Reference Architecture for Research Infrastructures?
• First step - State of the Architectures (5 EU ESFRI Social Science & Humanities)
• Towards a Reference Architecture
• Future for a Reference Architecture
Research Infrastructures

• Research is becoming increasingly interdisciplinary
• There is a multitude of research infrastructures (RIs)
• From 3 key stakeholder perspectives:
  – Researchers
  – Data archives
  – Funders
RIs - Researcher Perspective

• I have a cross-discipline research questions
• Interesting data is distributed
• Various policies for access, usage, etc.
  – Confusing and can take time
• Many differing metadata & data formats - does it all work with my tools?
• Domain specific terminology
• Lots of different portals
RIs - Data Archives Perspective

- Demands from different research infrastructures for access to metadata & data
- Requirements for researchers to deposit all scientific data
- Open (access) data
- Domain specific terminology
RIs - Funders Perspective

- Maximise high quality outputs from funding
- Minimise repetitive activities & redundant services
- Opening access to research outputs nationally & internationally
The Issues

• Roles & responsibility for stakeholders not clear.
• Misunderstandings are common
• Many research infrastructures in Europe
  – all in different phases (e.g. ESFRIs -> ERICs)
  – heterogeneous, plans, goals and communities supported
  – but at an abstract level often very similar (but this is implicit)
• 'Levels' at which an infrastructure functions
  – community, domain, special purpose, technical, other
DASISH acknowledges that these are normal problems across SSH disciplines for many stakeholders
What is a Reference Architecture?

• It is an instrument to compare, discuss & evolve research infrastructures
• Semantic alignment
• Used to define roles & responsibilities
  – Tiered accountability
• A roadmap for roadmaps
• Integration of existing reference models
What is a Reference Architecture?

• It should support heterogeneity between infrastructures (specificities)
• Identify commonalities and loci of interface
• Towards alignment of disparate architectures
• Identify points of departure & principles
Reference Architecture as mediation
State of the Architectures

• Bottom-up approach for comparison
• Looking for commonalities in available documentation for 5 SSH RIs + interviews
• Using a modified Zachman EA model matrix
  – 6 interrogatives: who, what, when, why, where, how.
  – Adapted levels of abstraction - policy/vision (scope), management (concepts & relationships), design (representations), specification (or demonstrators)
  – Depth & breadth
State of the Architectures

- Little commonality on broad interrogatives
- 'How' & 'What' addressed frequently but not 'When & Where'
- No uniform architectural view

<table>
<thead>
<tr>
<th></th>
<th>How</th>
<th>What</th>
<th>When</th>
<th>Where</th>
<th>Who</th>
<th>Why</th>
</tr>
</thead>
<tbody>
<tr>
<td>CESSDA</td>
<td>2</td>
<td>13</td>
<td></td>
<td>3</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>CLARIN</td>
<td>28</td>
<td>23</td>
<td>4</td>
<td>7</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>DARIAH</td>
<td>30</td>
<td>4</td>
<td>3</td>
<td></td>
<td></td>
<td>12</td>
</tr>
<tr>
<td>ESS</td>
<td>5</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SHARE</td>
<td>16</td>
<td>15</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
State of the Architectures

• All have in place advocacy, user consultation & collaboration layers
• Growing communities & knowledge sharing is central to all 5 infrastructures
• Infrastructures appear better placed to serve those already engaged with software tools
• Difficult to conclude commonalities & specificities
Table of Contents for a Reference Architecture

• Introduction
  – Audience
  – Scope, purpose & rationale
  – Principles
• Vocabulary
• Roles & responsibilities
• Model & Views
• Guidelines for implementation
Governance of the Reference Architecture

- Not a static reference model
- DASISH version - just a first iteration.
- Sustainable & usable
- Principles & model for governance
### Questions?

<table>
<thead>
<tr>
<th><strong>What?</strong></th>
<th>A set of common descriptions for responsibilities, roles, activities, instruments within an infrastructure.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>How?</strong></td>
<td>Assists the assessment and development of an infrastructure. The RA needs to be maintained.</td>
</tr>
<tr>
<td><strong>Where?</strong></td>
<td>The RA will be managed centrally but maintained de-centrally.</td>
</tr>
<tr>
<td><strong>Who?</strong></td>
<td>The RA is by- and for those that are responsible for research infrastructures. This can be on a strategic, functional or technical level, and during both the development and the operational phases.</td>
</tr>
<tr>
<td><strong>When?</strong></td>
<td>The RA should be used when changes to an infrastructure are planned, and it should be updated when the state of the art or common requirements of infrastructures change.</td>
</tr>
<tr>
<td><strong>Why?</strong></td>
<td>The RA improves communication between different stakeholders of different infrastructures. It provides a reference for defining responsibilities and functions within an infrastructure. It supports the identification of opportunities and threats within an existing infrastructure.</td>
</tr>
</tbody>
</table>
Thank you for your attention!