RODA Infrastructure for Data and Metadata

Cosmin Rentea
Data & Metadata

- RODA = Romanian Social Data Archive, Bucharest
- Currently: less than 100 Codebooks (Nesstar) and associated datasets (SPSS)
- Project Goals:
  - collect new data and metadata from Romanian public institutions and private organisations (doubling the archive contents after 2 years)
  - migrate existing data to DDI-Lifecycle
Data & Metadata

- RODA = Romanian Social Data Archive, Bucharest
- Currently: less than 100 codebooks (Nesstar) and associated datasets (SPSS)
- Goals:
  - collect new data and metadata from Romanian public institutions and private organisations (100% archive increase after 2 years)
  - migrate existing data to DDI-Lifecycle
Software Solution (1)

- Open-source solution
- Reusable extendable modules
- Data Model implementing DDI-Codebook and a subset of DDI-Lifecycle
- We are developing a multi-tier web application
  - Complemented by other dedicated applications (CRM, DMS, Search)
Software Solution (2)

- Security-aware Application:
  - users, roles
  - multiple authentication methods
  - authorization (ACL)
- Multiple persistence back-ends:
  - RDBMS
  - XML
- Indexing/Searching metadata
- CESSDA requirements for future integration: Shibboleth authentication, harvester access
Model-View-Controller pattern
Spring Framework

Data Access/Integration
- JDBC
- OXM
- ORM
- JMS
- Transactions

Web (MVC / Remoting)
- Web
- Servlet
- Portlet
- Struts

AOP
- Aspects
- Instrumentation

Core Container
- Beans
- Core
- Context
- Expression Language

Test
Database Schema
Hardware

- Storage Server
- Database Servers
- Web Server
- Email Server
- Backup Server

- Workstations
- Laptops
- Tablets
- Network equipments
- Printing equipments
Quality Assurance

- Build Manager
- Continuous Build System
- trac
- SVN
  etc.
Proposed Technologies

- Spring
- PostgreSQL
- Hibernate
- Ext JS
- R
- Apache Solr
- Shibboleth
- SugarCRM
- Alfresco
Closing Remarks

• RODA is dedicated to open-source software
  - Willingness to share, test, reuse, contribute
  - Modular application design

• Data Model (Java) for DDI standards can be a shared asset

• We can later share encountered problems and best practices in both software development and data migration processes