RODA

RODA's Open-Source Web Platform for DDI

Adrian Dușa

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

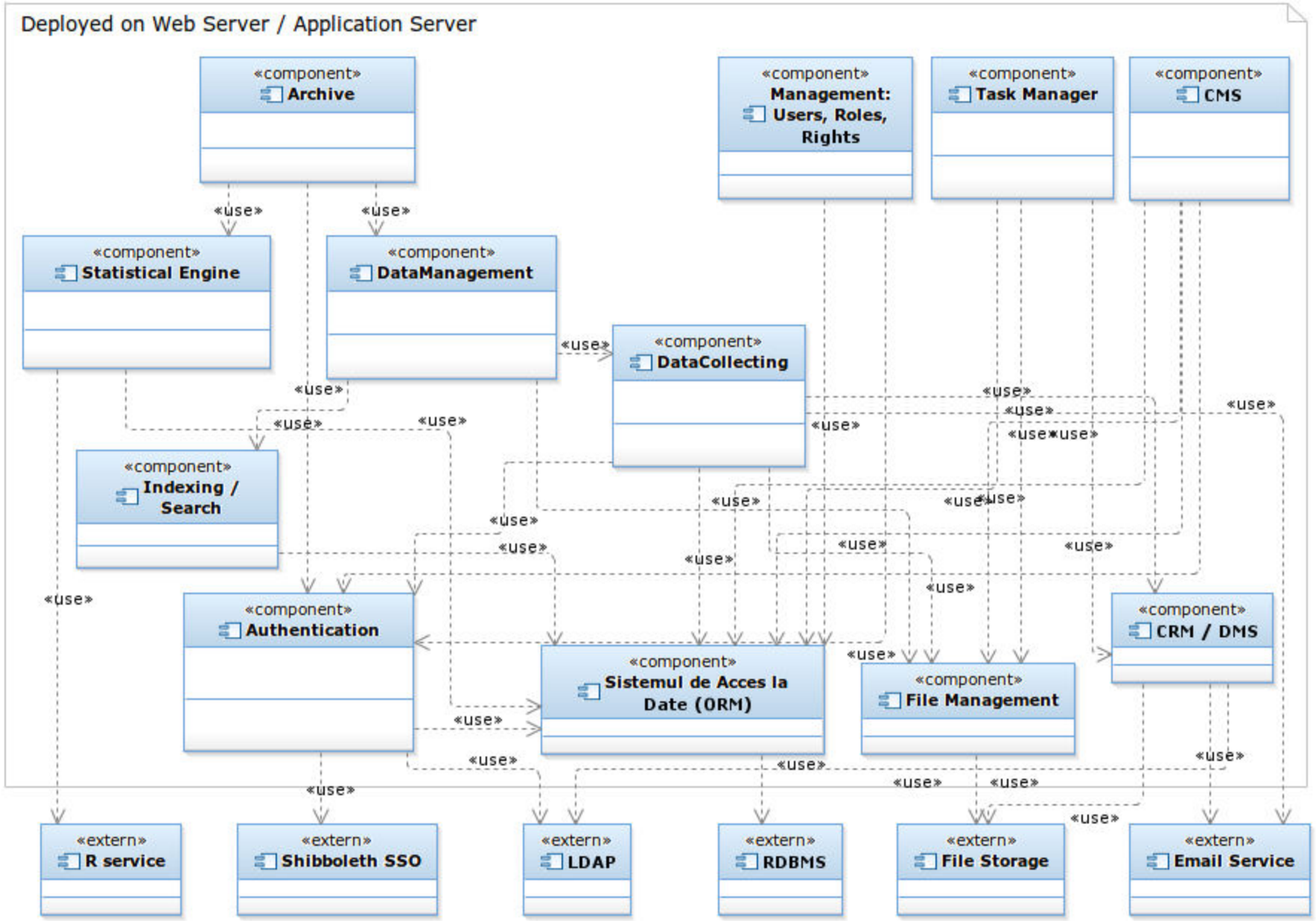
Software Solution

- 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
  - Server: Java-based; Clients: JavaScript;
  - Complemented by other dedicated applications (CRM, DMS, Search)

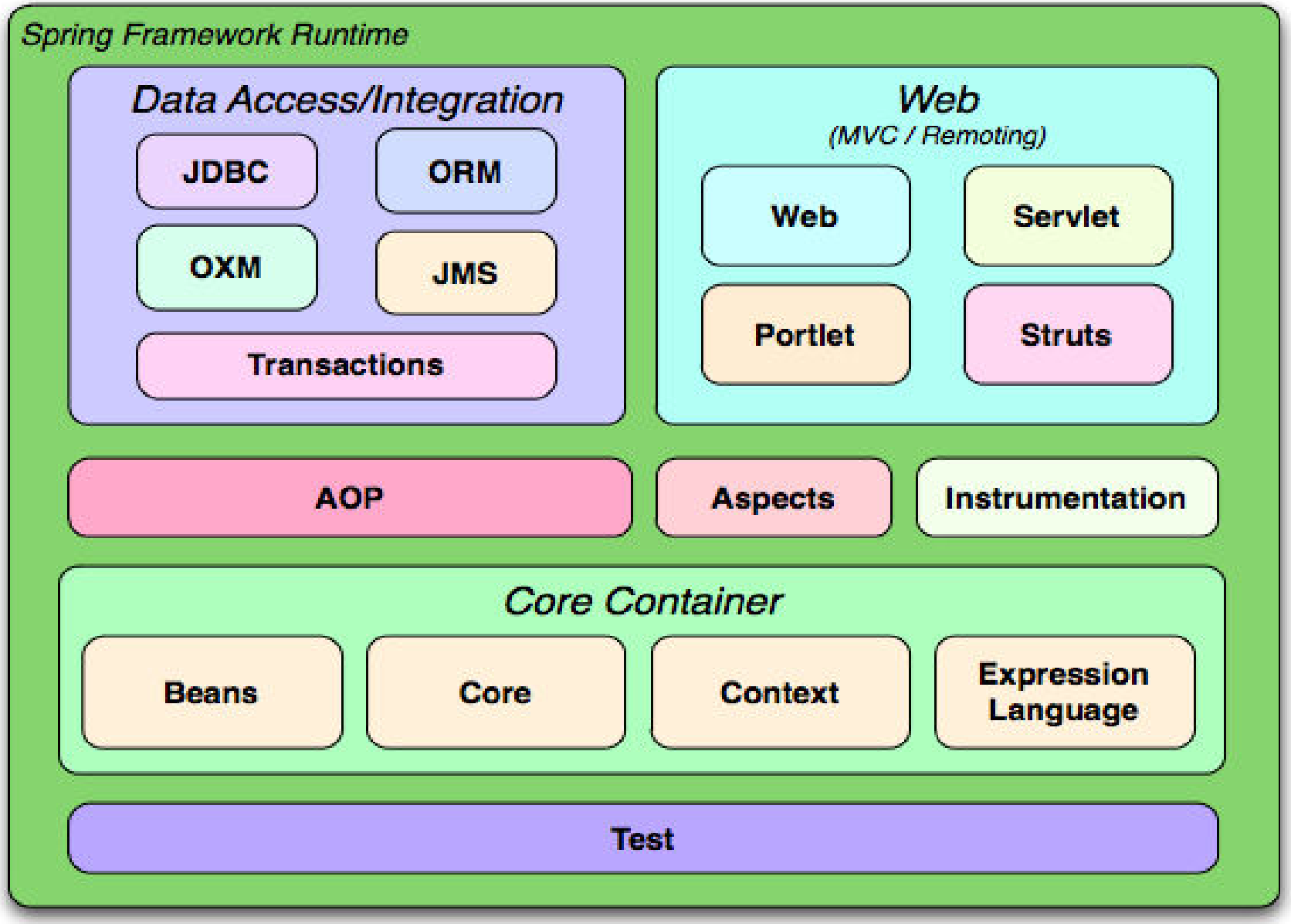
Software Solution

- Security-aware Application:
  - users, roles
  - multiple authentication methods
  - authorization (ACL)
  - possibility to use LDAP
- Multiple persistence back-ends:
  - RDBMS
  - XML
  - File Storage
- Indexing/Searching metadata
- CESSDA requirements for future integration: Shibboleth authentication, harvester access

UML Component Diagram



Spring Framework



Database

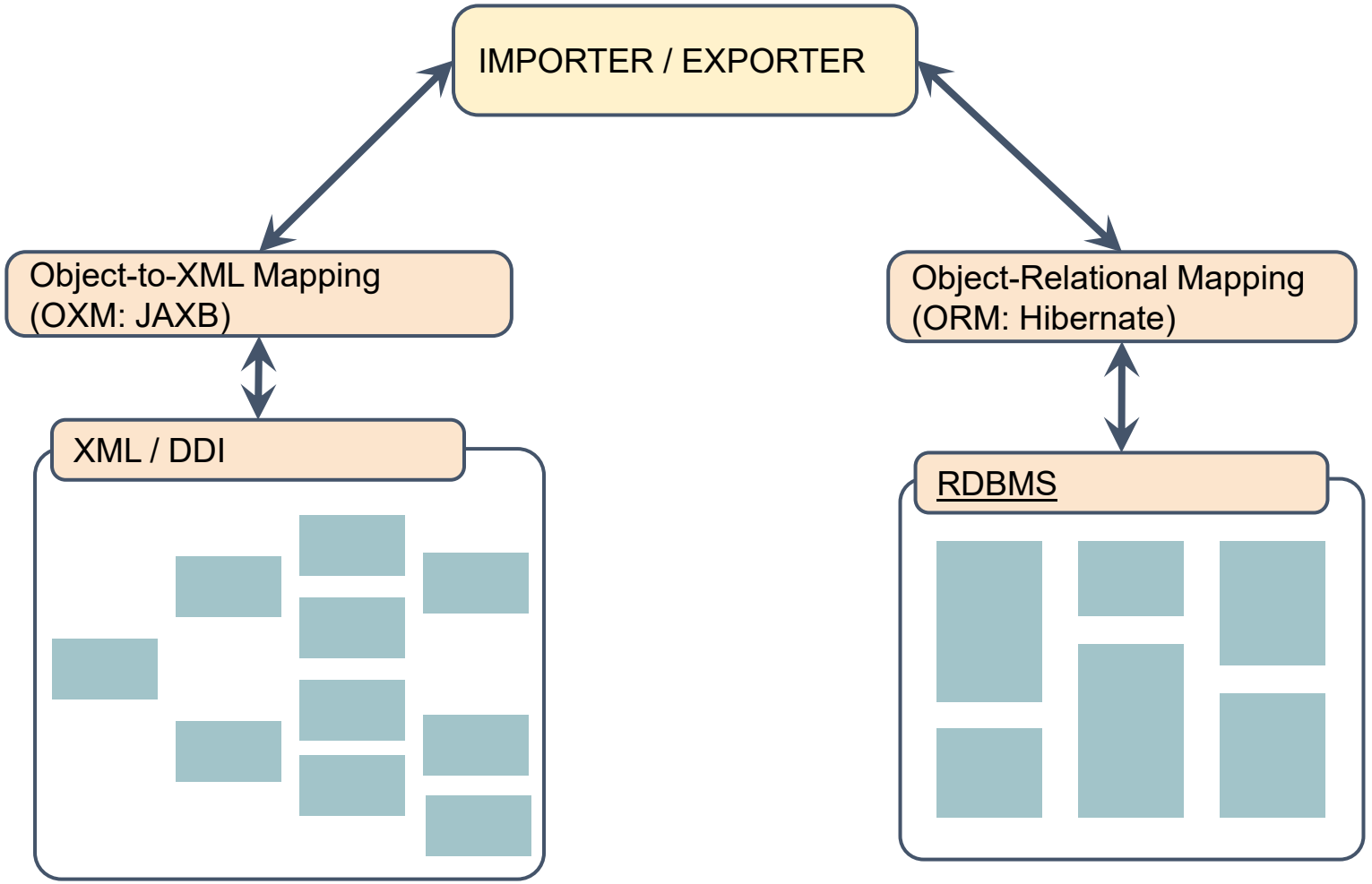
- Database Schemas for
  - RODA Schema:
    - Studies and related concepts (methodology …)
    - Catalogs, Datasets, Variables, Questionnaires
    - Persons & Organizations
    - Geographical Info
    - Thesaurus (ELSST), Keywords, etc.
  - DDI original data
  - CMS
  - ACL
  - Versioning information
- RDBMS choice: PostgreSQL

Components

- Web U.I. – based on ExtJS
- Task-Scheduling – using Spring
- Data Versioning – using Hibernate Envers
- Content repository – using JackRabbit for
  - Codebook-related files
  - CMS
- Statistical module – using R

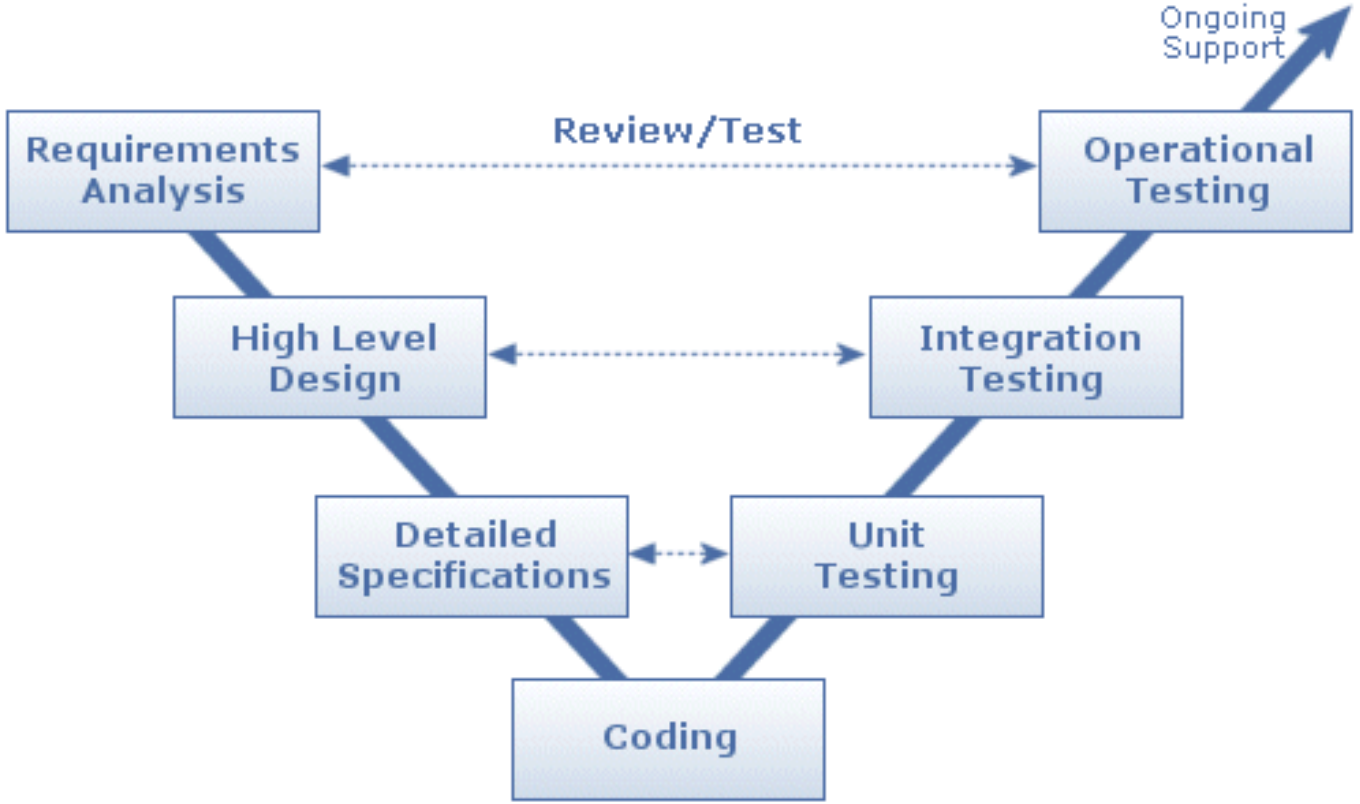
Components

- Importer & Exporter – DDI data, other files (CSV, SQL, SPSS)



Quality Assurance

- Tests
  - Unit Tests (JUnit)
  - Integration Tests
    - Spring Testing Framework
    - Selenium
- Build manager
  - Maven
- Continuous Build System
  - Jenkins
- Version-Control
  - SVN private repository
    (for the Archive's data)
  - Github for code : https://github.com/cosminrentea/roda



Hardware used by the software platform

- Storage Server
- Database Server(s)
- Web Server
- Email Server
- **Backup** Server
- **LDAP** server

- Workstations, Laptops, Tablets

Technologies

Ext JS

spring source

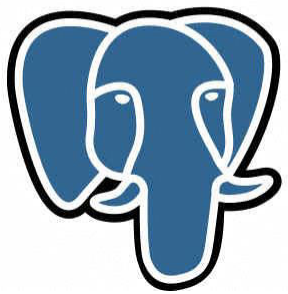
SUGARCRM

HIBERNATE

Apache Solr

Apache Jackrabbit™

PostgreSQL



Shibboleth

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
- Many Components reusable for Social Sciences online platforms:
  - CMS, DMS, Task-Scheduling, DB Versioning, Importer/Exporter
- We can later share the encountered issues & best practices in both software development and data migration process