RZ Information System Meeting

Ing. José A. Mejía Villar M.Sc.
jmejia@awi.de

Computing Center of the Alfred Wegener Institute for Polar and Marine Research

29. November 2011

Contents

1. Fedora Commons Repository
2. Federico
3. Federico's Live Demo

1. Fedora Commons

1. Fedora Commons

1.1. What is Fedora Commons?

1.2. Key Features vs Disadvantages

1.3. Web Service Interfaces

1.4. Framework Services

1.5. Who uses Fedora Commons in Germany?

1.1 What is Fedora Commons?

- Fedora stands for *Flexible Extensible Digital Object Repository*.
- Fedora is a general-purpose, open-source digital object repository system.
- Java based conceptual framework using a set of abstractions about digital information to provide the basis for software systems that can manage digital information.
- The Fedora software distributed by Duraspace (http://www.duraspace.org) is available from http://fedora-commons.org under the terms of the Apache License, version 2.0.

1.2 Key Features [1/3]

- Store all types of content and its metadata
- Scale to millions of objects
- Access to data via Web APIs (REST/SOAP)
- Provides RDF based Resource Index search
- Rebuilder Utility (for disaster recovery and data migration)
- The entire repository can be rebuilt from the digital object and content files.

1.2 Key Features [2/3]

- Content Model Architecture (define "types" of objects by their content)
- Many storage options (database and file systems)
- JMS messaging provider (your apps can "listen" to repository events)
- OAI-PMH Provider Service

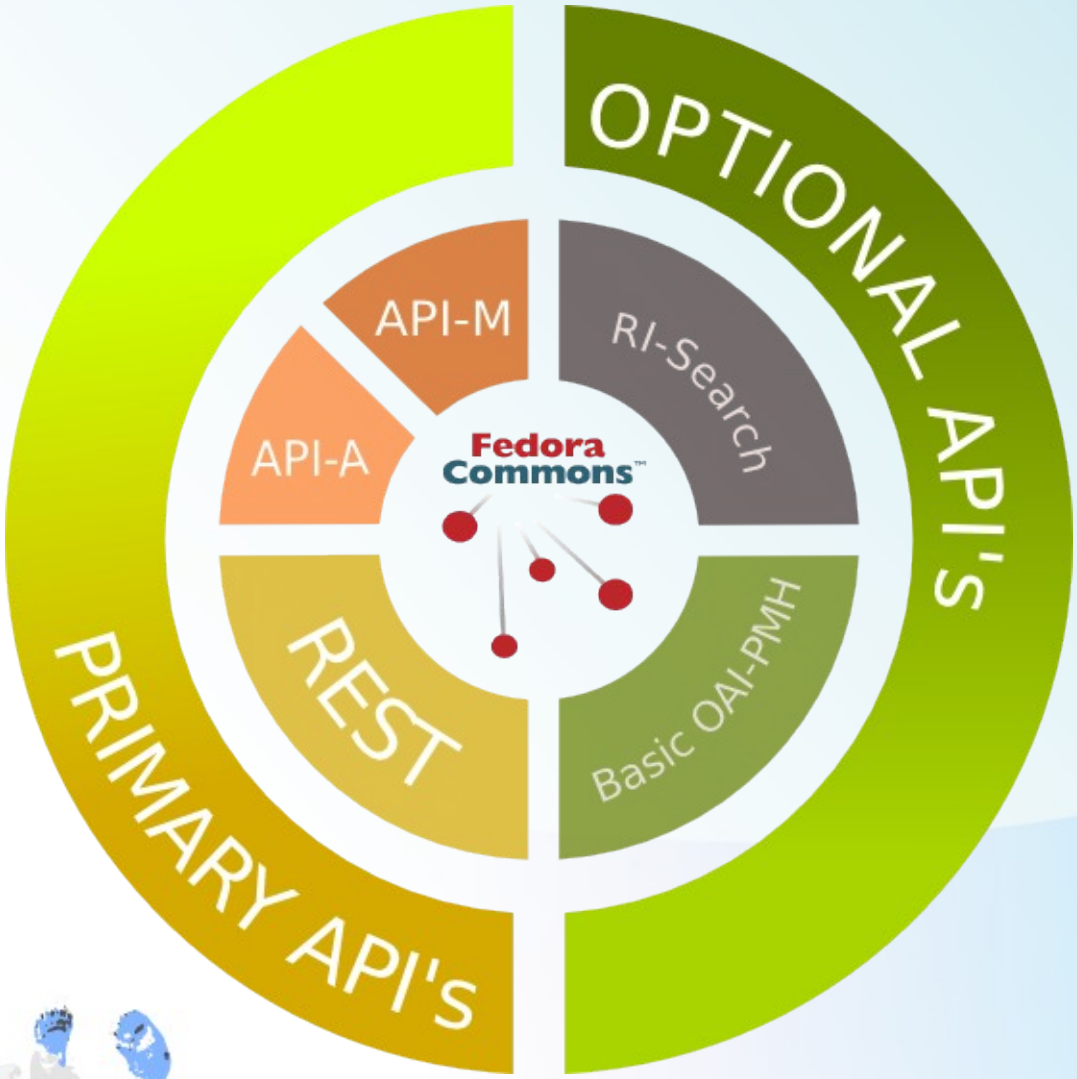
1.2 Disadvantages [3/3]

- Front-end Adaptation

  https://wiki.duraspace.org/display/DEV/Fedora+Tools
- Object Store Scalability Strategy

  https://wiki.duraspace.org/display/AKUBRA/Akubra+Project

1.3 Web Service Interface

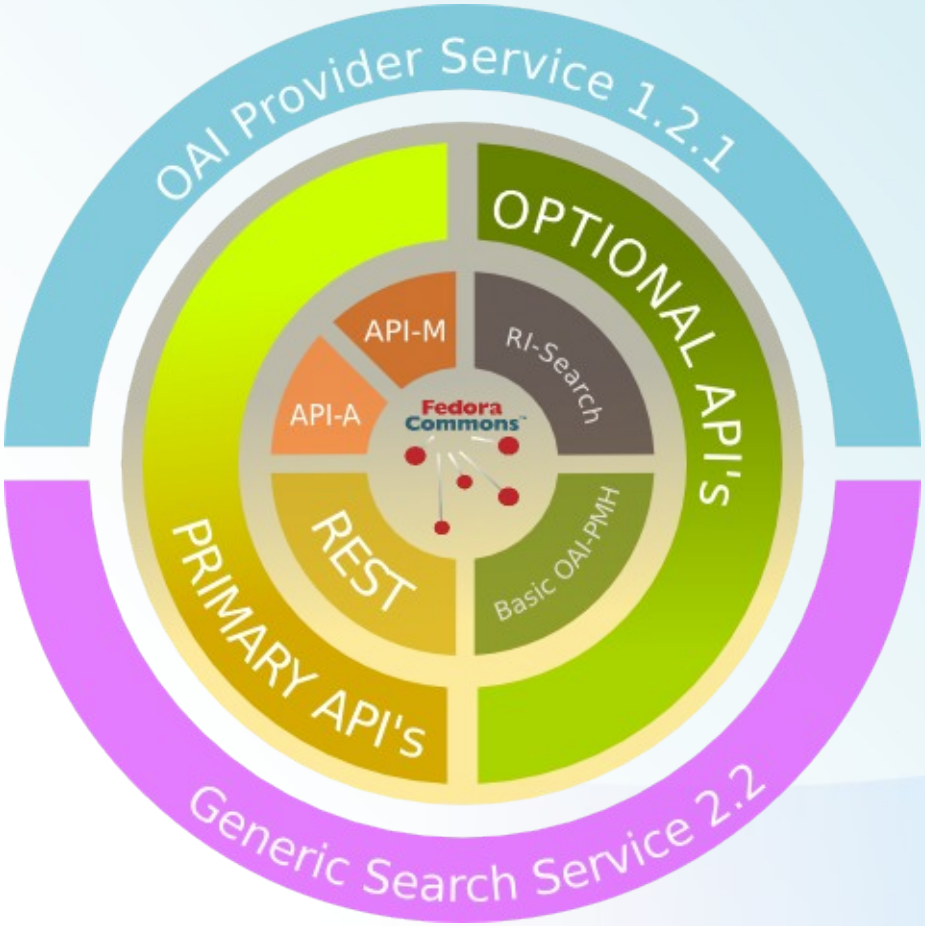


- **Primary API's**

  Allow the creation, reading, modification, and deletion of Fedora digital objects.

- **Optional API's**
  - Basic OAI-PMH
  - RI-Search

1.4 Framework Services



- Generic Search Service
- OAI Provider Service

1.5 Who uses Fedora Commons in .de?

1- eSciDoc - The Open Source e-Research Environment
https://www.escidoc.org

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

MAX-PLANCK-GESELLSCHAFT

2- Fedora-IRODS Integration als Grid Repository (in development)
http://www.wissgrid.de/publikationen/deliverables/wp3_de.html

WissGrid

NIEDERSÄCHSISCHE STAATS- UND UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN

WissGrid

2. Federico

2.1. What is Federico?

2.2. System Requirements

2.3. Content Model

2.4. Architecture

2.5. Future Plans

2.1 What is Federico?

- Fedora-Enabled Repository with Cocoon
- AJAX-based frontend for a C3Grid local repository of metadata
- Transparent Integration of Fedora with the Framework Services GSearch and OAI Provider
- Developed in the scope of the work package #3, Long-term Preservation of Digital Archives of Wissgrid, sponsored by the German Federal Ministry of Education and Research

2.2 System Requirements [1/2]

Hardware

- PC with a 1 gigahertz (GHz) processor or faster and network card
- 2 GB RAM
- 800 MB free disk space for the installation

Software

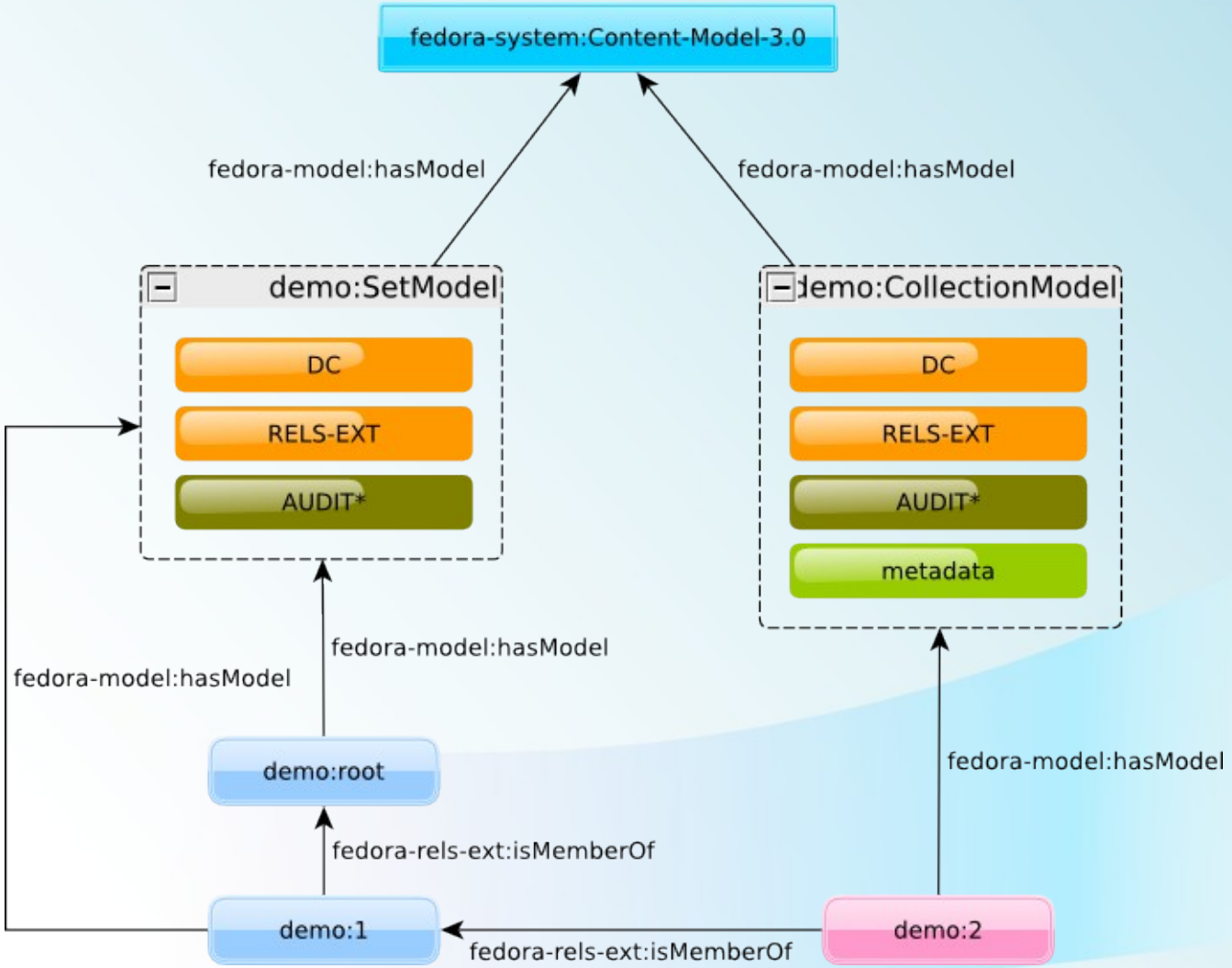
- Linux Distribution with X Window System
- Java JDK 1.6
- 3 MySQL Databases for Fedora Commons, Fedora OAI Provider, and openID accounts

2.2 System Requirements [2/2]

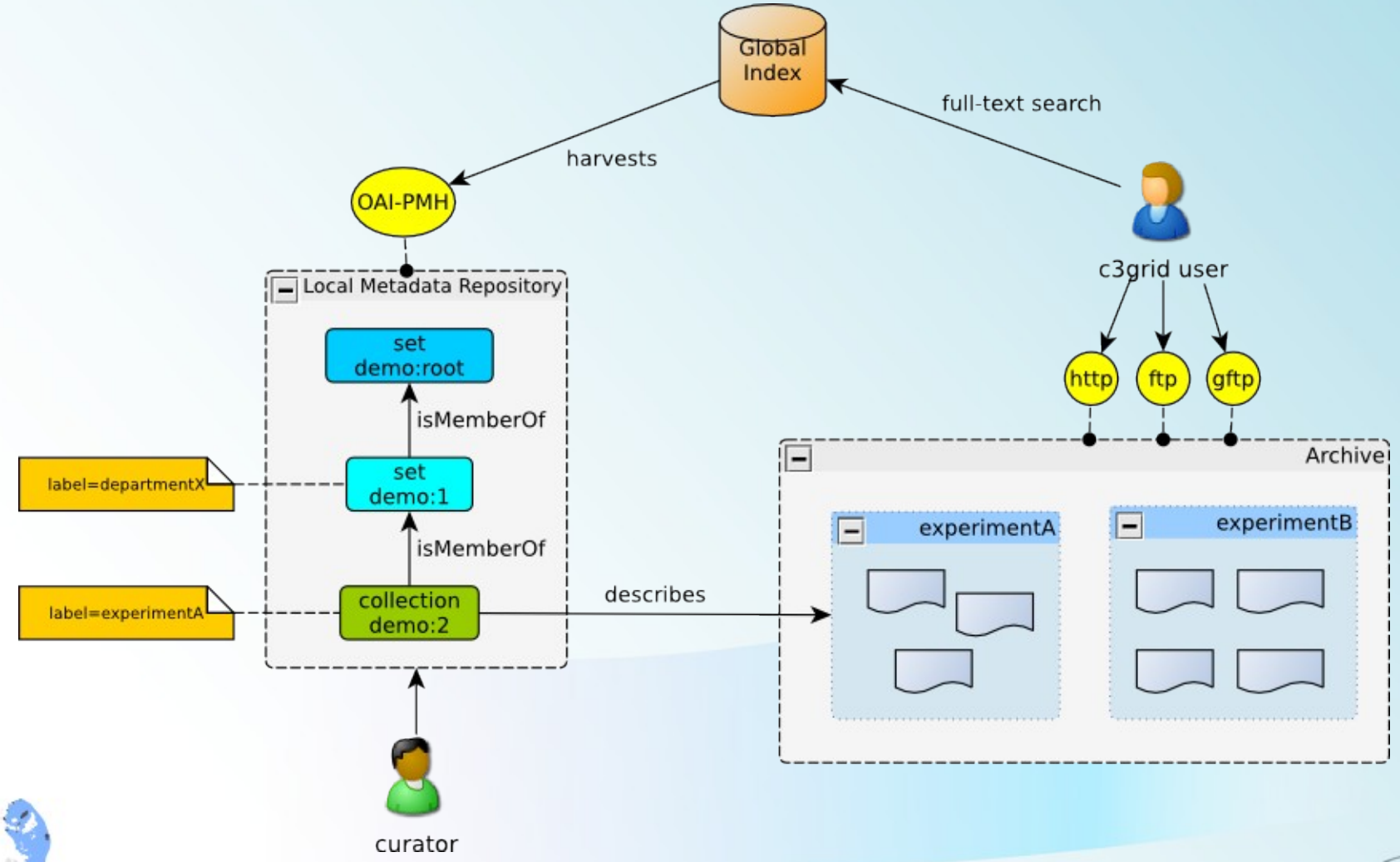
User

- PC with graphical interface and network card
- Keyboard and mouse
- Browser (preferably Mozilla Firefox) with Javascript enabled

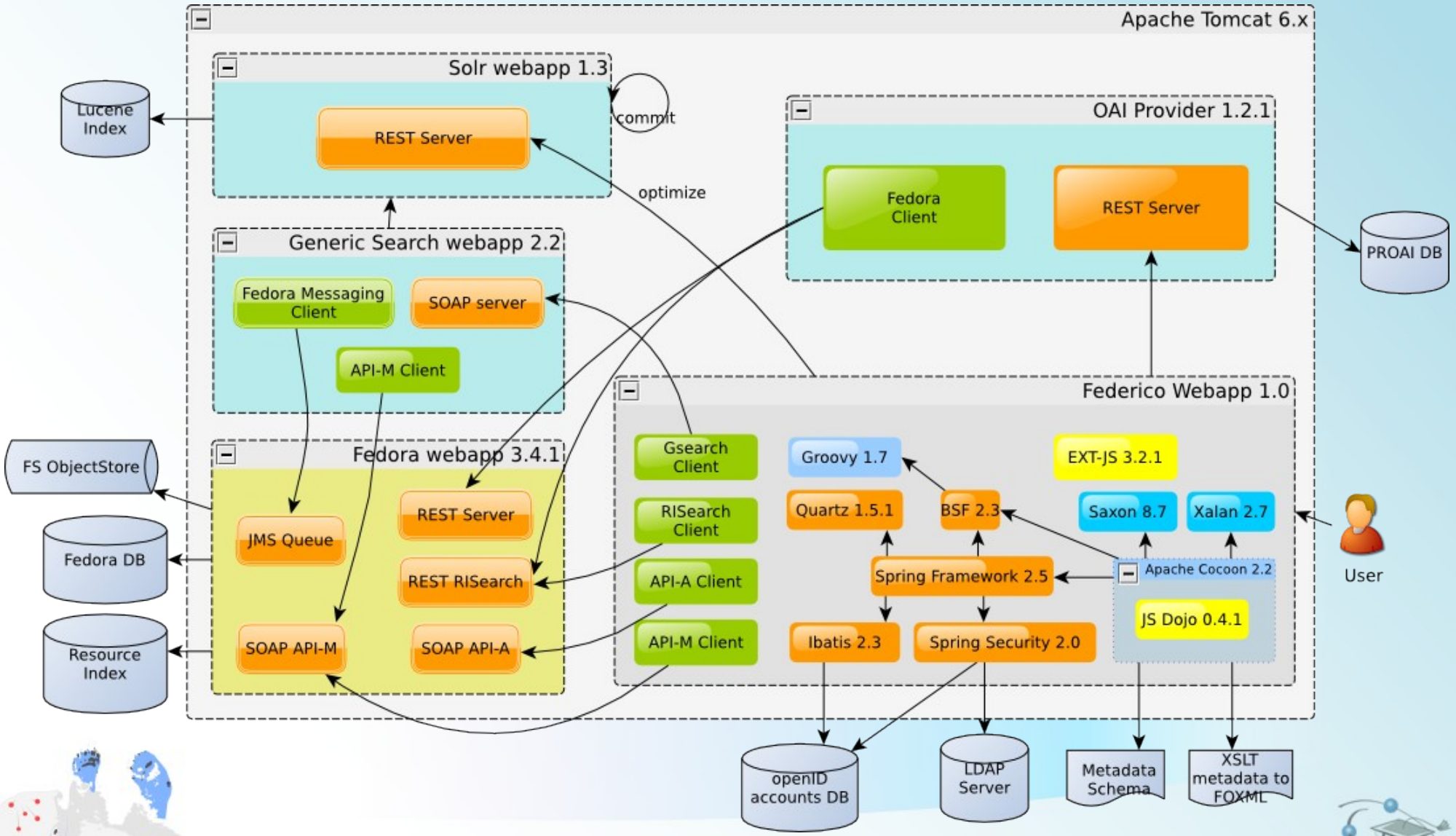
2.3 Content Model [1/2]



2.3 Content Model [2/2]



2.4 Architecture



2.5 Future Plans

- Hibernate Persistence Layer
  - To support other database manager systems: Postgresql
- Reduction of code
  - One client for API-M, API-A, RI-Search
- Easier configuration and maintainability
- Upgrade of Spring 2.5 to 3.0
- Support of spatial queries

3. Federico's Live Demo

3.1. User Interface

3.2. Authentication

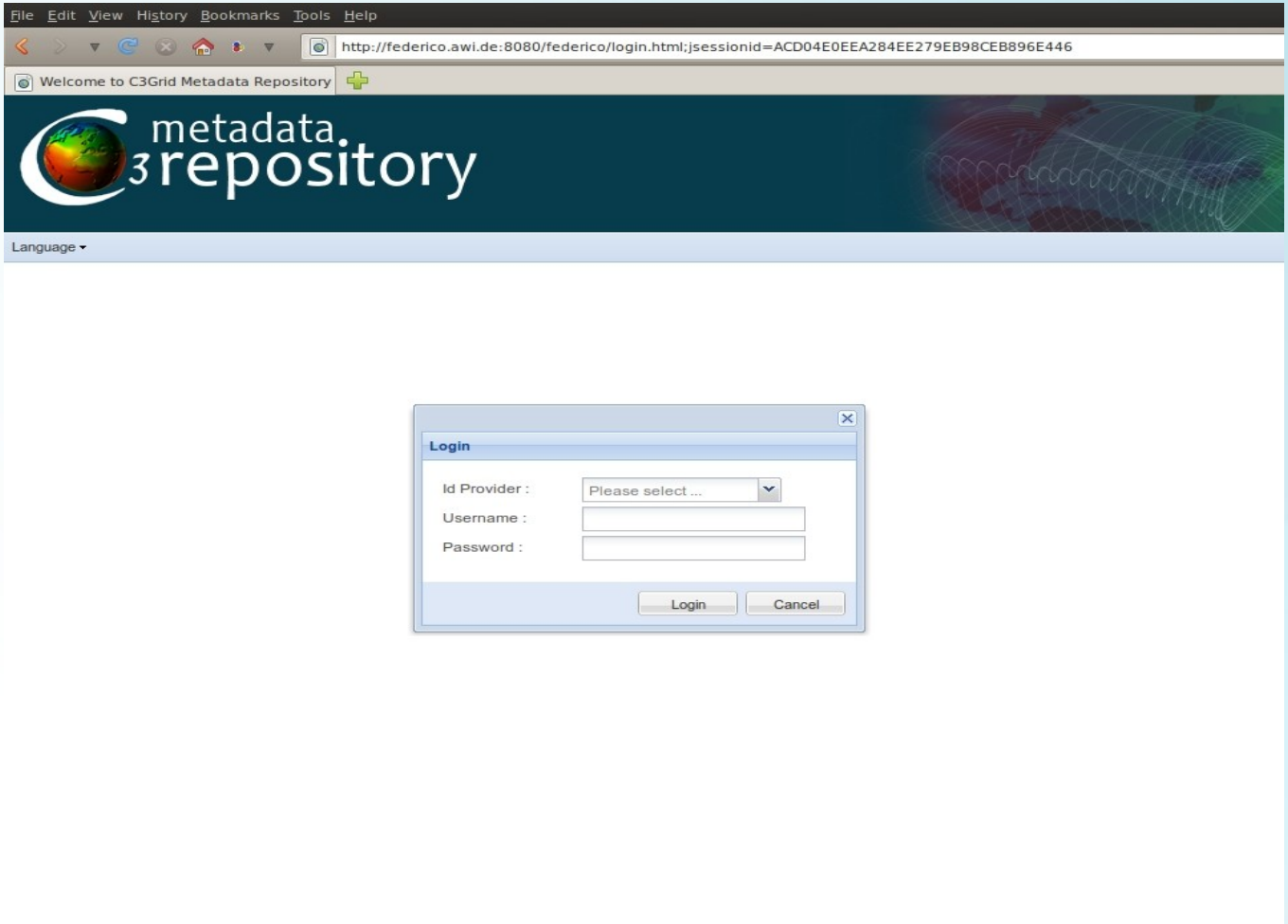
3.3. Ingest Collections

3.4. Full-text Search

3.5. OAI-PMH

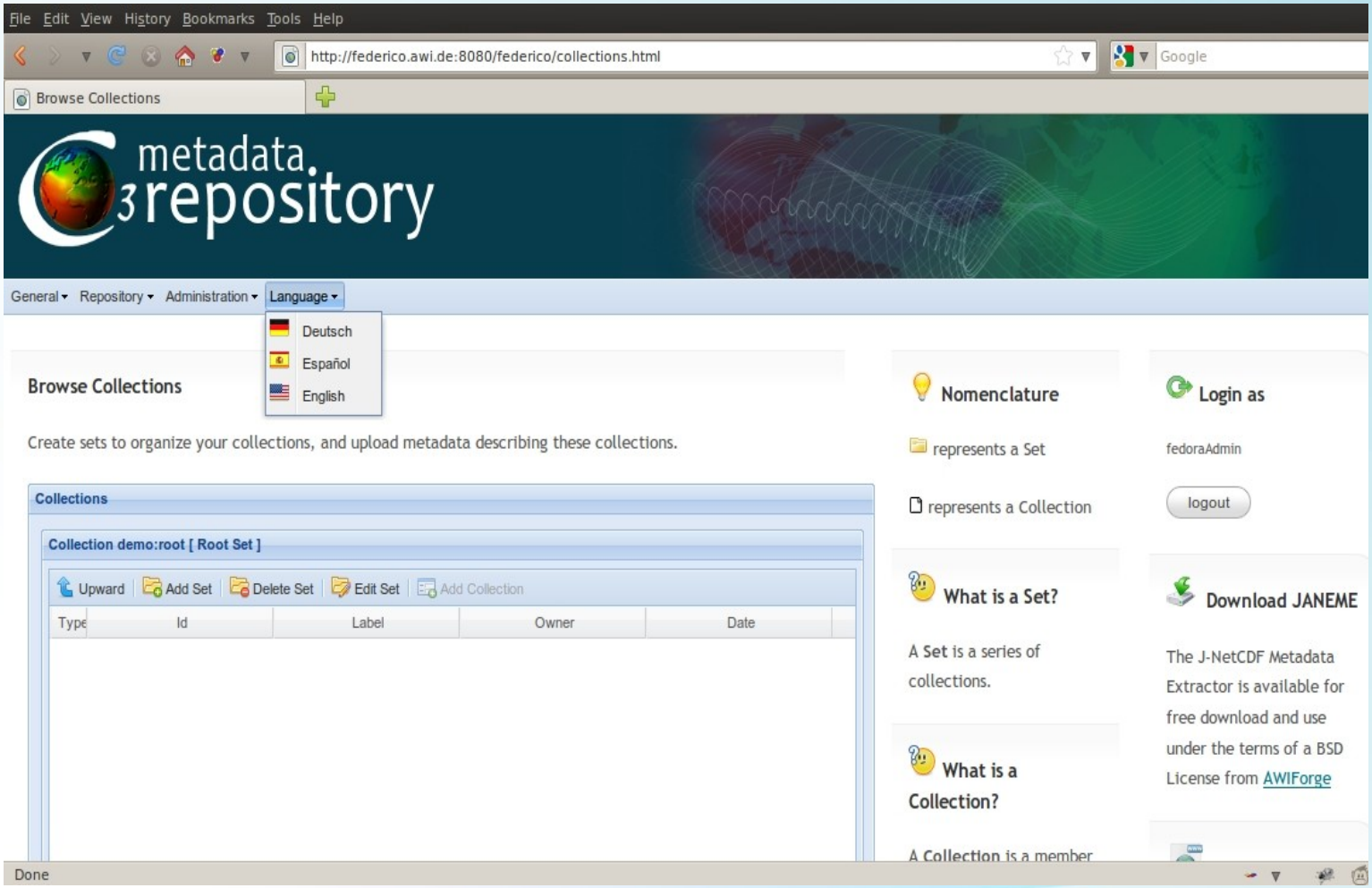
3. Federico's Live Demo: Screenshots

Login Form



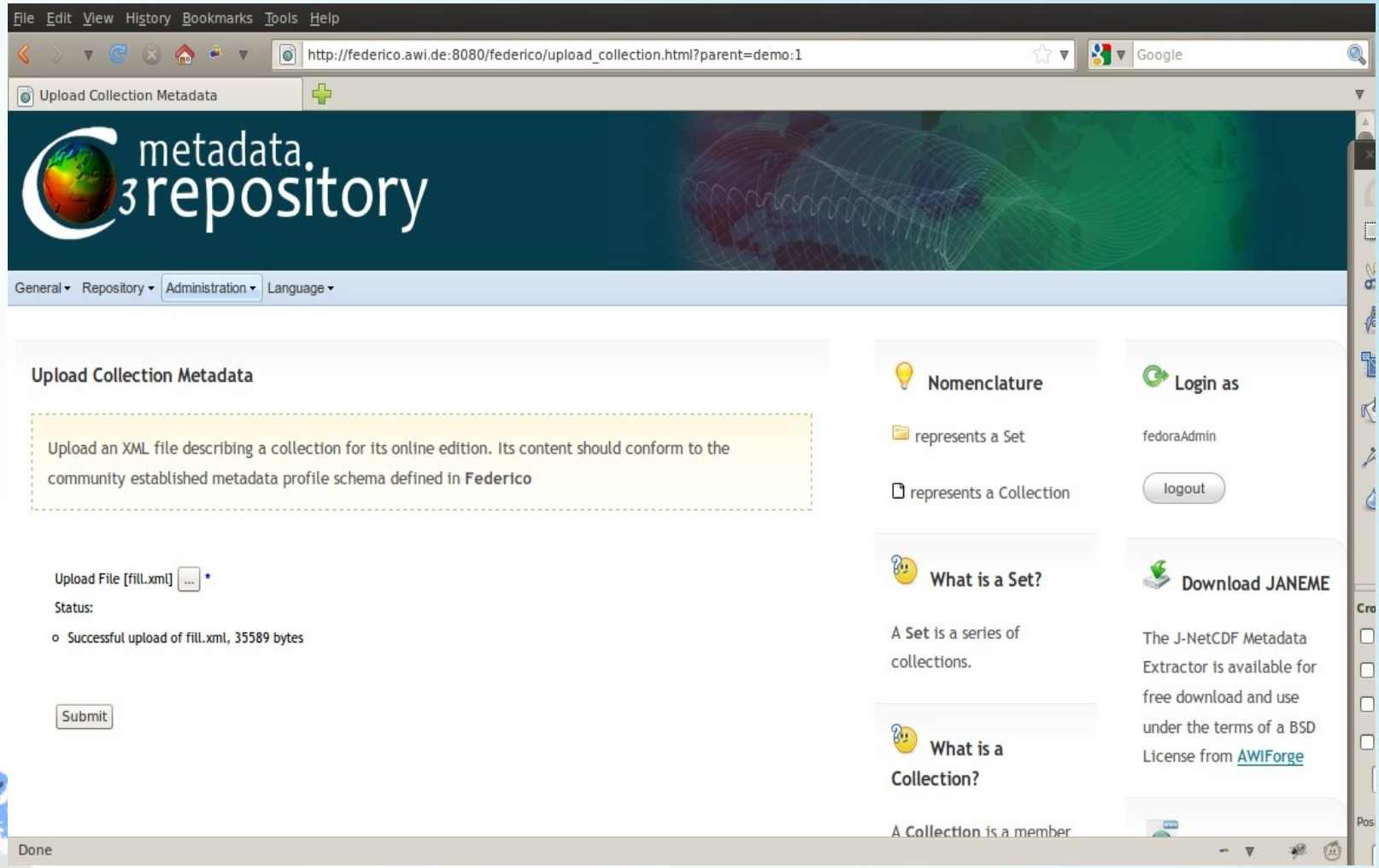
3. Federico's Live Demo: Screenshots

Browse Collections Form



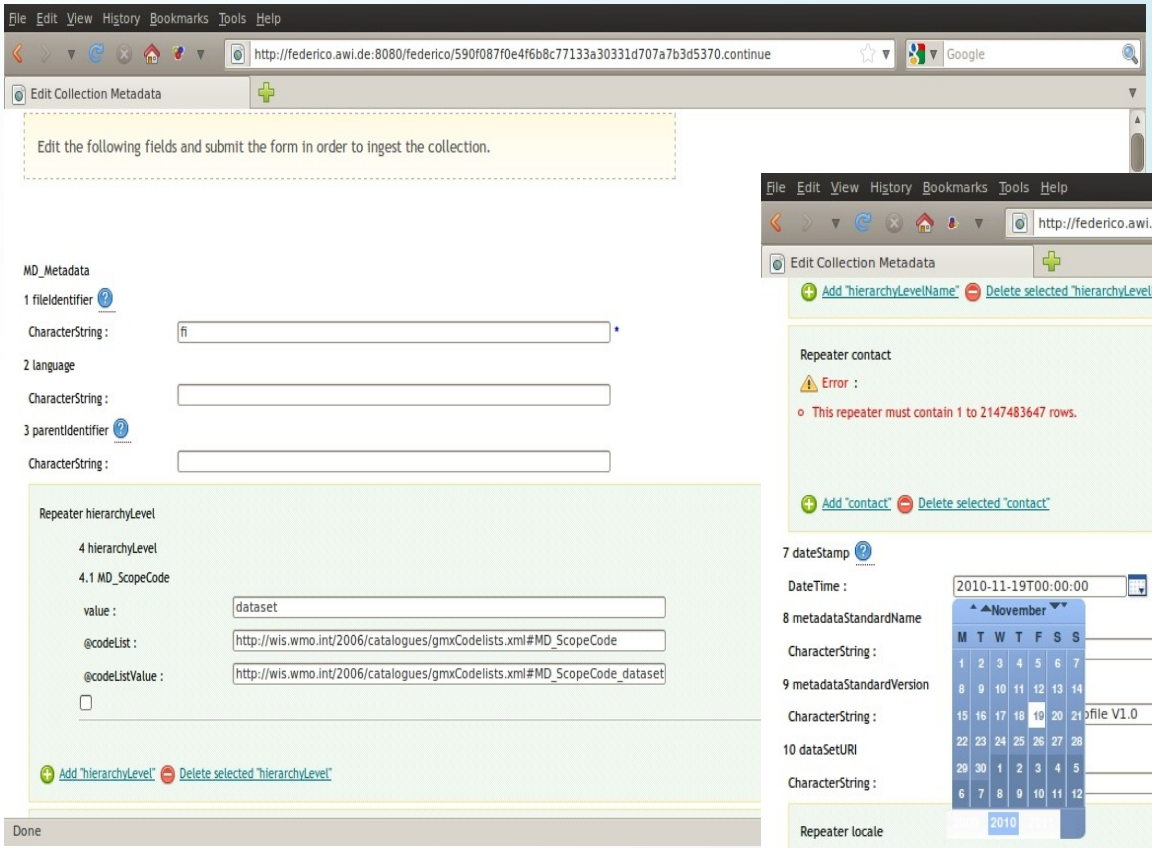
3. Federico's Live Demo: Screenshots

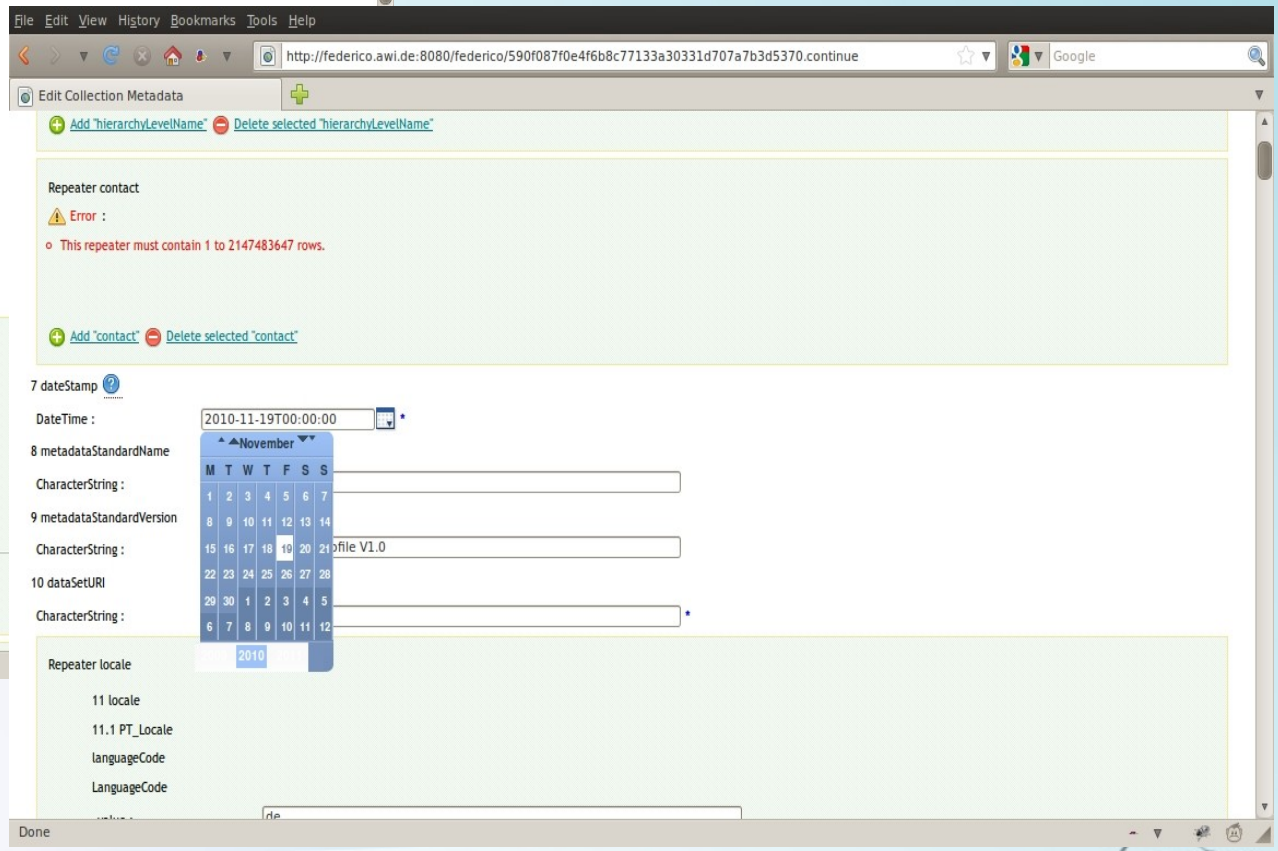
Metadata Upload



3. Federico's Live Demo: Screenshots

Metadata Edition





Summary

- Fedora as repository for digital information in research environment
  - Well defined API's
  - Content Model Architecture for the definition of “types” of objects
  - Harvesting through OAI-PMH
- Knowledge of XML is crucial
- Difficult UI implementation