

RZ Information System Meeting



federico

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

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

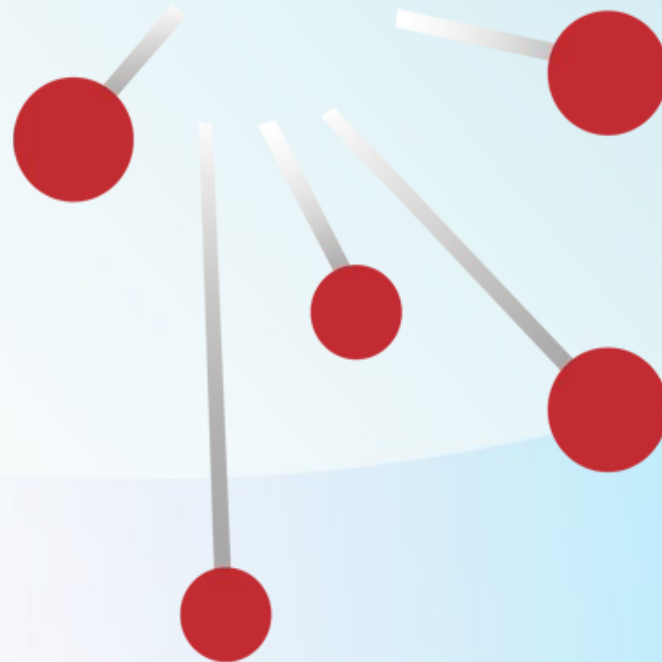
29. November 2011

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



1. Fedora Commons

Fedora
CommonsTM



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

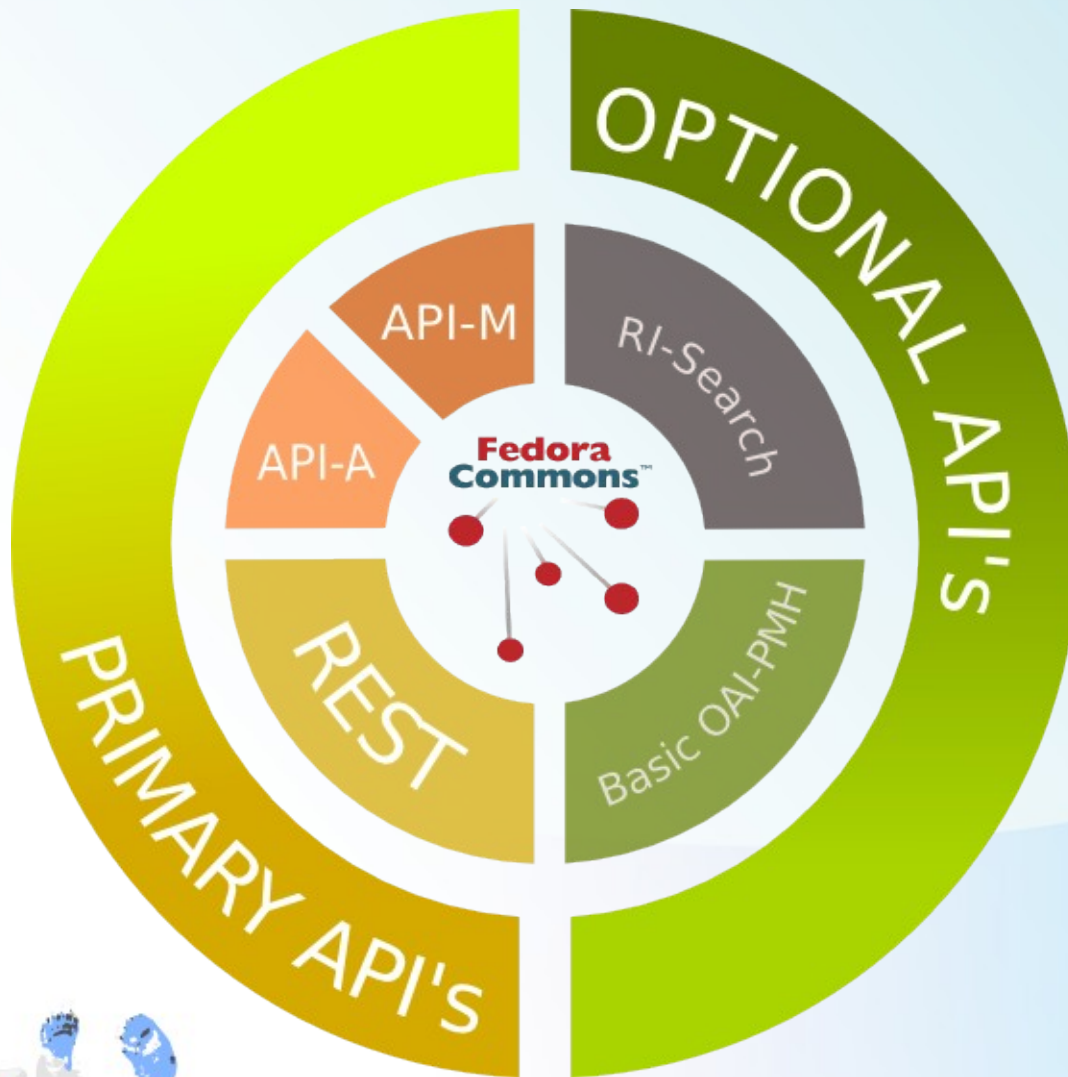
<https://wiki.duraspace.org/display/DEV/Fedora+Tools>

- Object Store Scalability Strategy

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



1.3 Web Service Interface



- **Primary API's**

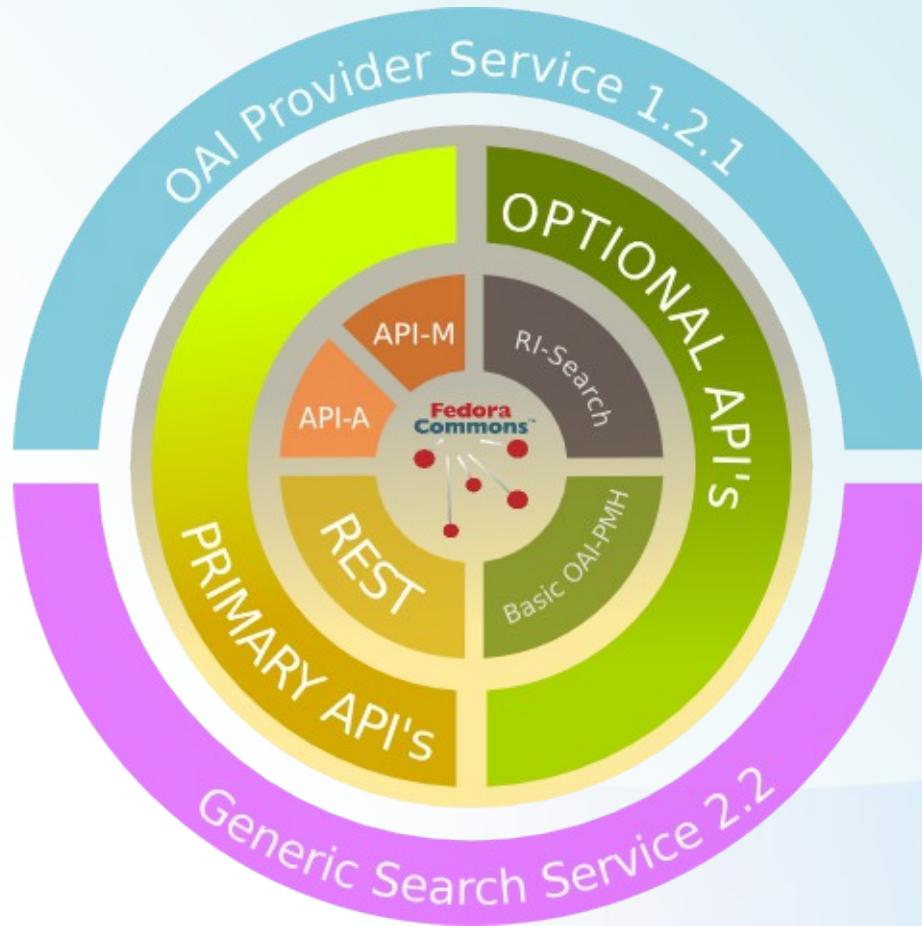
Allow the creation, reading, modification, and deletion of Fedora digital objects.

- **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>



MAX-PLANCK-GESELLSCHAFT

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





- 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



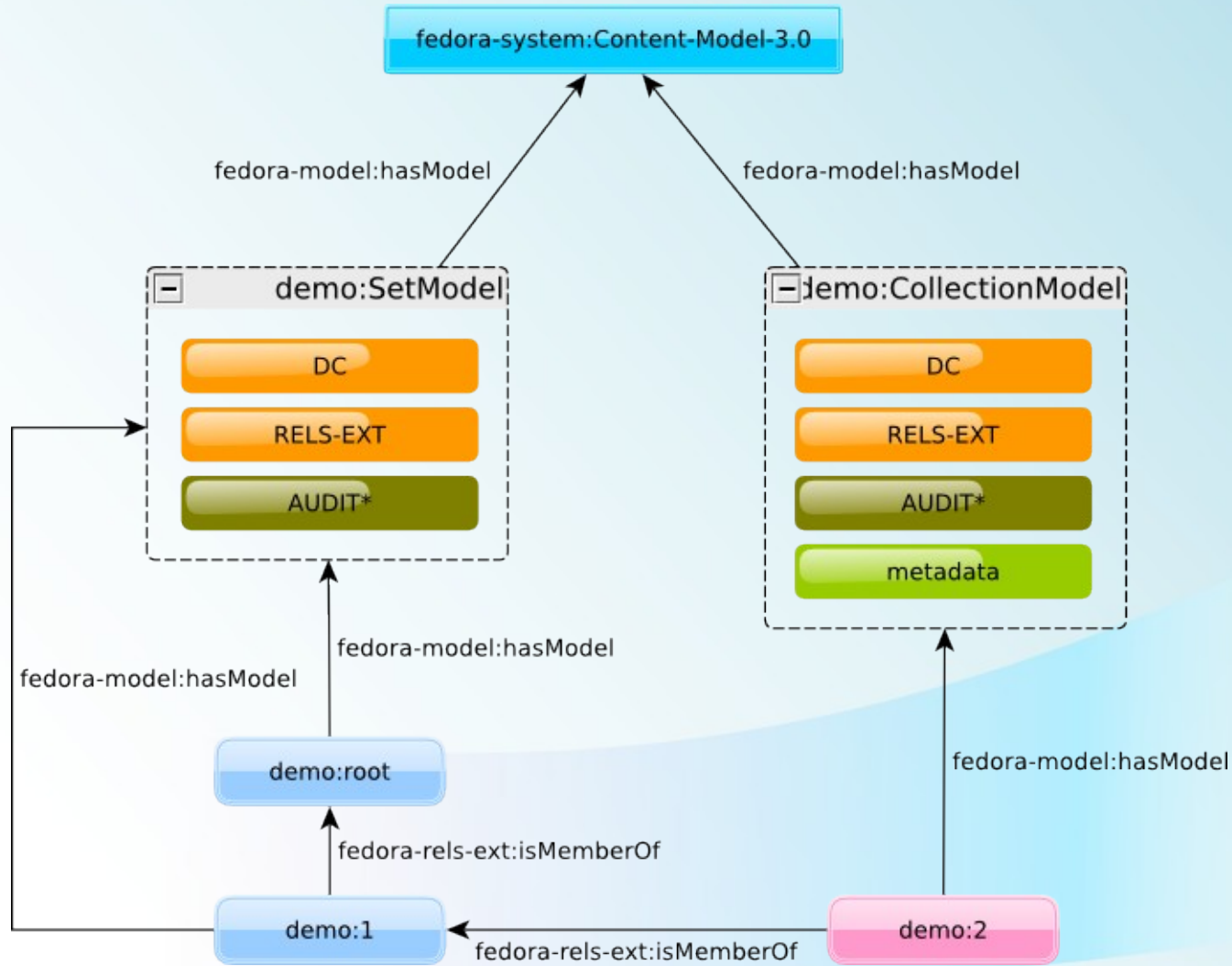
2.2 System Requirements [2/2]

User

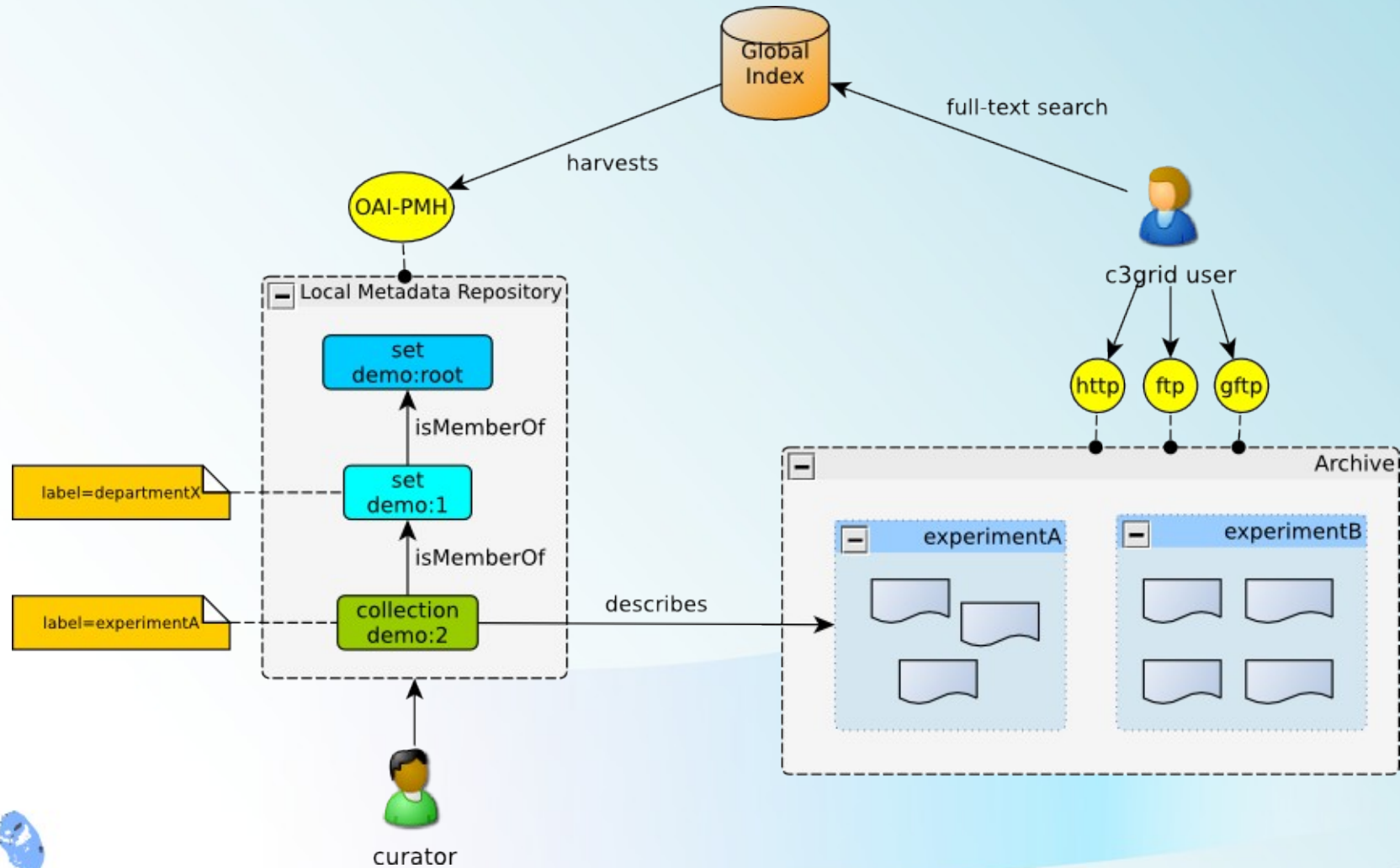
- PC with graphical interface and network card
- Keyboard and mouse
- Browser (preferably Mozilla Firefox) with Javascript enabled



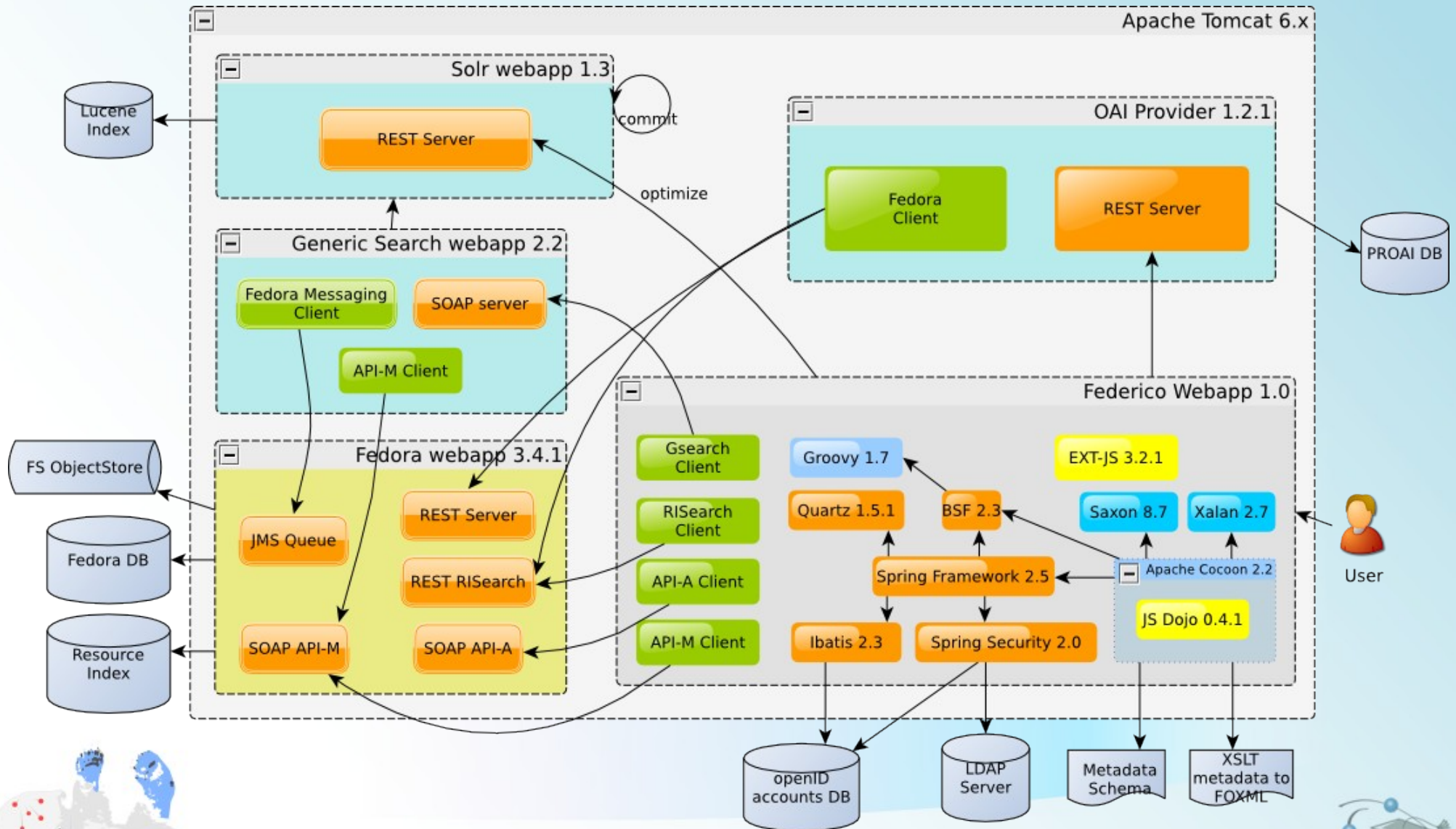
2.3 Content Model [1/2]



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

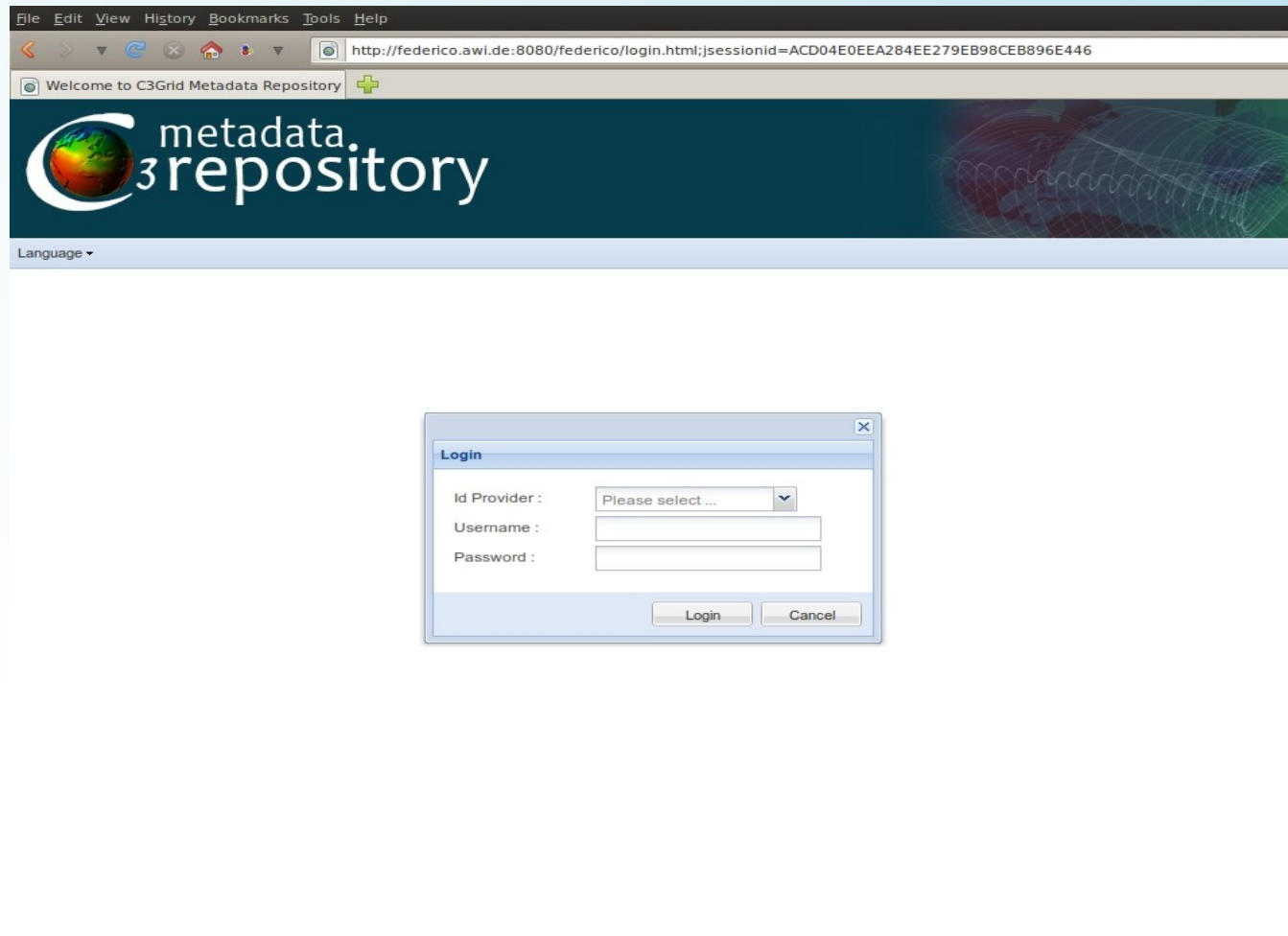


- 3.1. User Interface
- 3.2. Authentication
- 3.3. Ingest Collections
- 3.4. Full-text Search
- 3.5. OAI-PMH



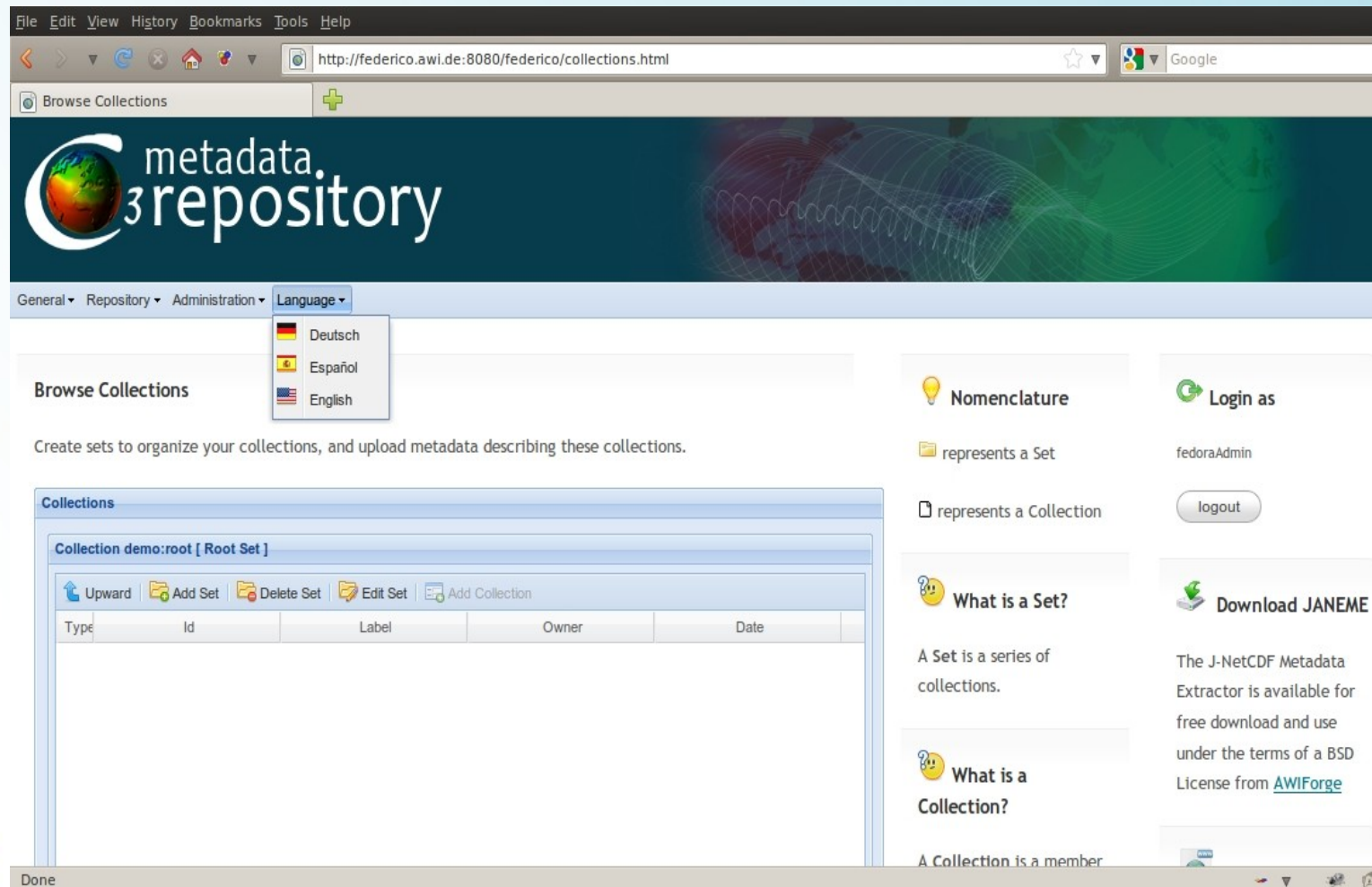
3. Federico's Live Demo: Screenshots

Login Form



3. Federico's Live Demo: Screenshots

Browse Collections Form



File Edit View History Bookmarks Tools Help

http://federico.awi.de:8080/federico/collections.html

Browse Collections

metadata repository

General Repository Administration Language

Deutsch
Español
English

Browse Collections

Create sets to organize your collections, and upload metadata describing these collections.

Collections

Collection demo:root [Root Set]

Upward Add Set Delete Set Edit Set Add Collection

Type	Id	Label	Owner	Date
------	----	-------	-------	------

Done

Nomenclature

represents a Set

represents a Collection

What is a Set?

A Set is a series of collections.

What is a Collection?

A Collection is a member

Login as

fedoraAdmin

logout

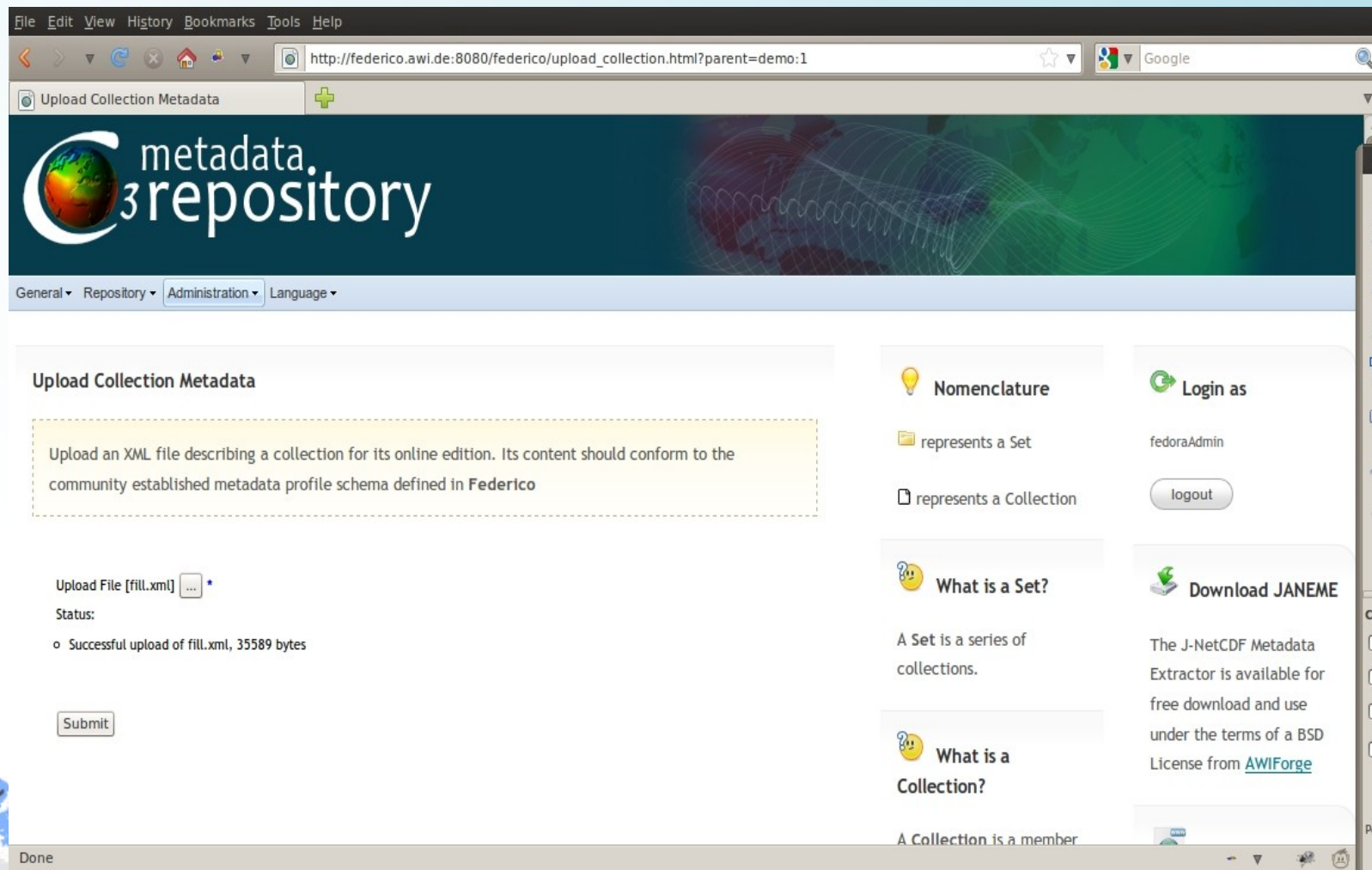
Download JANEME

The J-NetCDF Metadata Extractor is available for free download and use under the terms of a BSD License from [AWIForge](#)



3. Federico's Live Demo: Screenshots

Metadata Upload

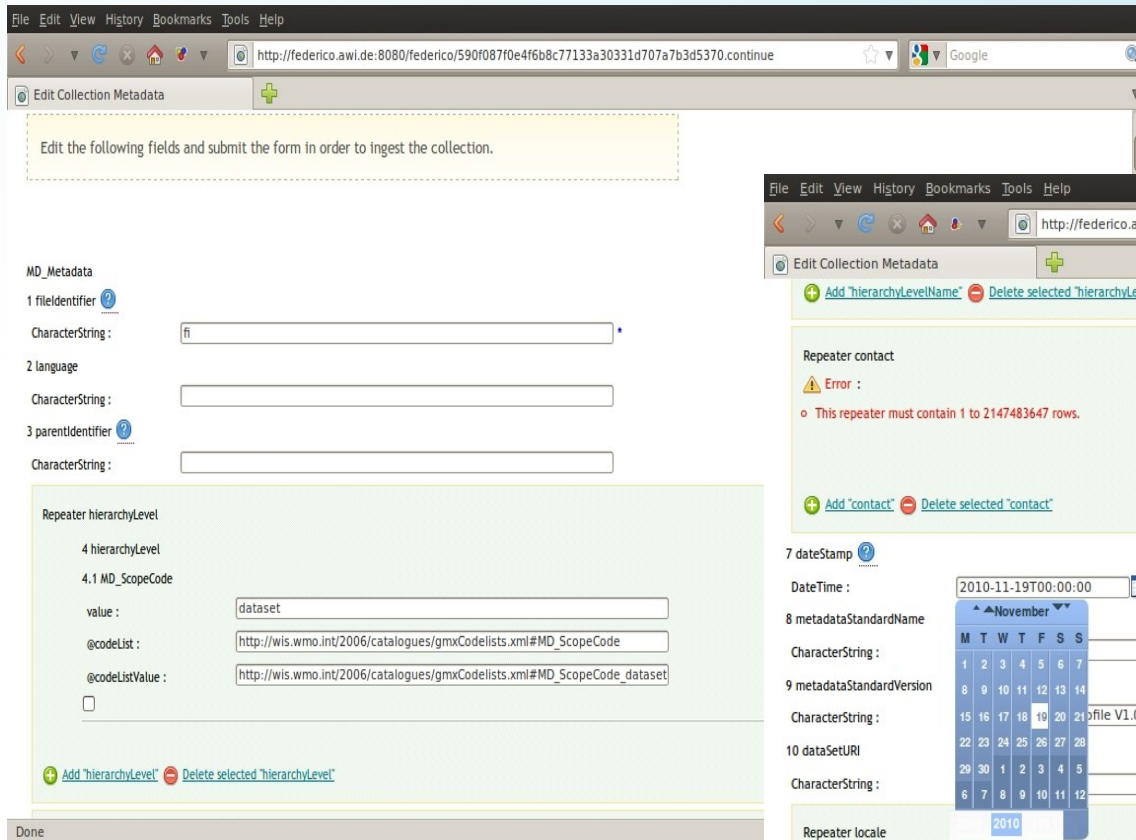


The screenshot shows a web browser window displaying the 'Upload Collection Metadata' page of the Federico Metadata Repository. The browser's address bar shows the URL: `http://federico.awi.de:8080/federico/upload_collection.html?parent=demo:1`. The page header features the 'metadata repository' logo and navigation menus for 'General', 'Repository', 'Administration', and 'Language'. The main content area is titled 'Upload Collection Metadata' and includes a yellow dashed box with instructions: 'Upload an XML file describing a collection for its online edition. Its content should conform to the community established metadata profile schema defined in Federico'. Below this, there is a file upload section with a text input field containing 'fill.xml', a 'Submit' button, and a status message: 'Successful upload of fill.xml, 35589 bytes'. To the right, there are several informational panels: 'Nomenclature' (explaining 'Set' and 'Collection' icons), 'Login as' (showing 'fedoraAdmin' and a 'logout' button), 'What is a Set?' (defining a set as a series of collections), and 'Download JANEME' (providing information about the J-NetCDF Metadata Extractor). The browser's status bar at the bottom shows 'Done'.



3. Federico's Live Demo: Screenshots

Metadata Edition



File Edit View History Bookmarks Tools Help

http://federico.awi.de:8080/federico/590f087f0e4f6b8c77133a30331d707a7b3d5370.continue

Edit Collection Metadata

Edit the following fields and submit the form in order to ingest the collection.

MD_Metadata

1 fileIdentifier

CharacterString : fi

2 language

CharacterString :

3 parentIdentifier

CharacterString :

Repeater hierarchyLevel

4 hierarchyLevel

4.1 MD_ScopeCode

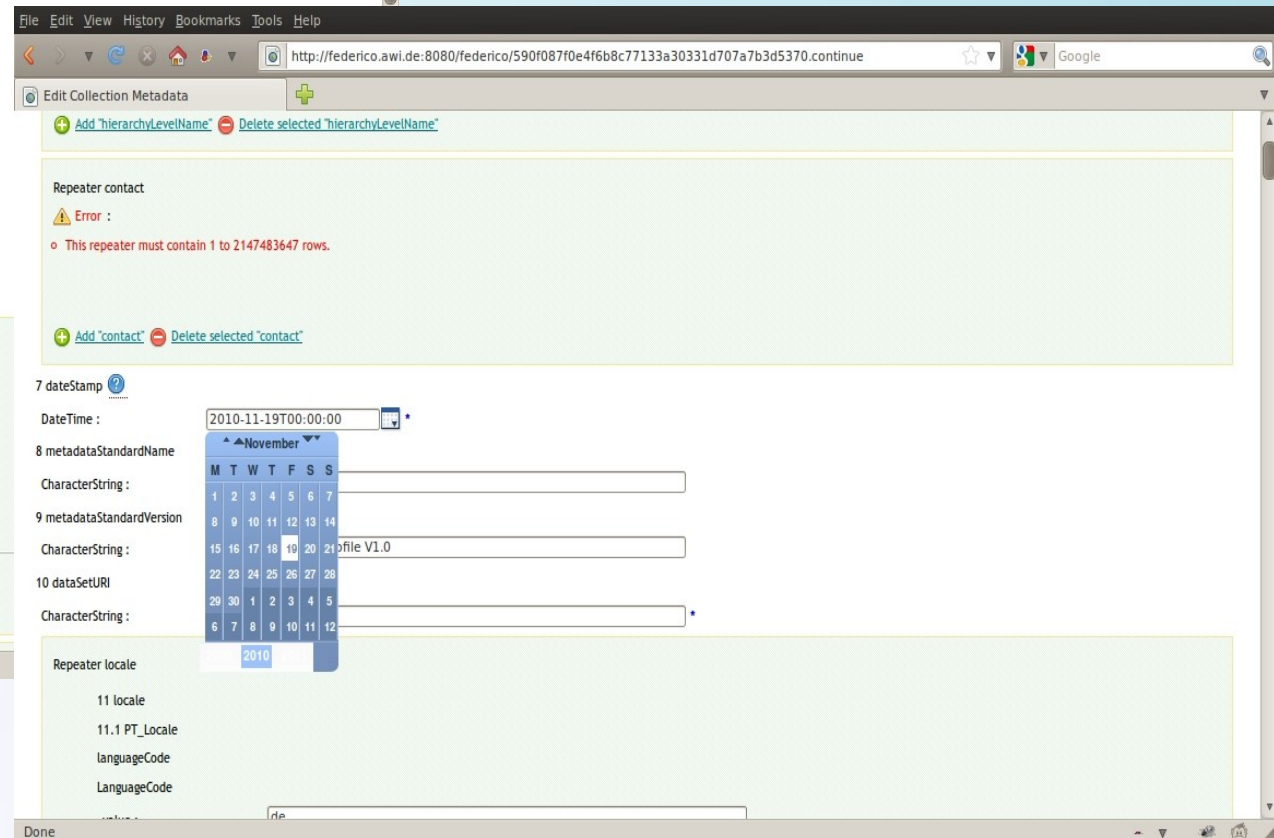
value : dataset

@codeList : http://wis.wmo.int/2006/catalogues/gmxCodeLists.xml#MD_ScopeCode

@codeListValue : http://wis.wmo.int/2006/catalogues/gmxCodeLists.xml#MD_ScopeCode_dataset

+ Add "hierarchyLevel" - Delete selected "hierarchyLevel"

Done



File Edit View History Bookmarks Tools Help

http://federico.awi.de:8080/federico/590f087f0e4f6b8c77133a30331d707a7b3d5370.continue

Edit Collection Metadata

+ Add "hierarchyLevelName" - Delete selected "hierarchyLevelName"

Repeater contact

Error :

This repeater must contain 1 to 2147483647 rows.

+ Add "contact" - Delete selected "contact"

7 dateStamp

DateTime : 2010-11-19T00:00:00

8 metadataStandardName

CharacterString :

9 metadataStandardVersion

CharacterString : file V1.0

10 dataSetURI

CharacterString :

Repeater locale

11 locale

11.1 PT_Locale

languageCode

LanguageCode

Done



- 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

