Participation and functional development @ MPDL

Visit from CAS
29th march 2010
Ulla Tschida
Scenarios to be addressed

Scenarios

Services

MPDL Projects

Publication Management

eSciDoc.PubMan

Linguistic texts – enhancing PubMan

JusCMS

Scholarly Workbench

eSciDoc.Faces

eSciDoc.VIRR

WALS

AWOB

Digitization lifecycle

XML Workflow
MPDL as partner for Max Planck Institutes

- **Human Sciences Section**
  - Bibliotheca Hertziana, Rome
  - Kunsthistorisches Institut, Florenz
  - MPI for Comparative and International Private Law
  - MPI for European Legal History
  - MPI for Evolutionary Anthropology
  - MPI for History of Science
  - MPI for Human Development
  - MPI for Intellectual Property, Competition and Tax Law
  - MPI for Psycholinguistics

- **Biology & Medicine Section**
  - MPI for Chemical Ecology
  - MPI for Infection Biology
  - MPI of Molecular Plant Physiology

- **Chemistry, Physics & Technology Section**
  - MPI for Biogeochemistry
  - MPI for Extraterrestrial Physics
  - MPI for Gravitational Physics
  - MPI of Colloids and Interfaces

- **Active project relations to 16 MPI's**
  - Projects develop specific solutions and generic components
    - Specific solutions for the institutes
    - Generic components for eSciDoc infrastructure
Positioning the MPDL within the MPS: MPDL Projects

- Strengthen the links to the MPIs
- Annual budget allotted by the President for
  - collaborative projects between institutes and the MPDL, which
  - benefit more than one MPI and are
  - of relevance to researchers.

- 7 projects are running so far
- 2 proposals in preparation
- Currently more project ideas than capacity
Many Services – one Infrastructure

Publication Management
- Solution PubMan

Scholarly Workbench
- Solution Faces
- Solution ViRR

“Publications”

“Images”

“Digitized Text and Image”
How to shape?
Dimensions and Measures

- Strategic goals
  - MPS boards (steering committee) and MPDL divisions
  - Projects

- Users
  - What do they need and why?
  - Where do they start from?
  - What are their priorities? Chances for compromise?
  - How to address conflicting interests for features?

- -> Internal coordination
  - Are we on time? Do we concentrate on the right things?
  - Where can we improve efficiency?
  - Where do we have communication gaps?
  - Where do we have to extend competence?
  - How do we address changing priorities?
The Role of Service Management

Scenarios

Idea

Expectation

Services

???

MPDL Projects

Solution

Satisfaction

Understanding the domain and the user concerns
Our Tasks

- Develop domain knowledge
- Translate ideas and expectations into specification
- Coordinate necessary activities
- Support end-users during development and production
- Set-up and maintain user communities
Our Tools

- CoLab Mediawiki
- Support and User Mailing lists
- Blogs
- Events (Workshops, institute visits, PubMan Days)
- Internal Meetings (=> GUI, => Dev)
- Issue Tracking
- Roadmaps
User Communities

Publications – trans-disciplinary, generic needs (PubMan)
• Pilot Group since 2004 (8 Institutes)
• Early Adopter (3 Institutes)
• Scientists, Librarians, local IT, Research Coordination
• Prioritized needs – vital interest

Research Data – research-specific needs
• User groups= researcher group with local-specific needs
• Generalize => other research questions to be addressed
• MPI Human Development => Material Science (NIMS) => MPI Ornithology (FACES)
• MPI European Legal History => MPI Art history (ViRR)
eSciDoc.PubMan – Publications

Current status:

• 6000 items released
• 5 Early Adopters
• 3 productive users
• Parallel Migration from eDoc

Already now comparable with top repositories (cf. Repository Matrix)
PhD Todoroki, Shin-ichi
National Institute for Materials Science, Optronic Materials Center

Current Position
Senior Researcher (National Institute for Materials Science)

Current Position
Faculty Member (Optronic Materials Center)

Researcher ID
urn:co:one:persons1008

Researcher Homepages
Researcher Blog

Research Fields
Computer Sciences; Free, Libre and Open Source Software; Information Science; Digital Libraries; Engineering; Materials Engineering; Ceramic Engineering; Optical Engineering; Informatics; Library Science; Digital Libraries; Physics; Materials Science

External references
WorldCat
Search for Todoroki, Shin-ichi

Google Scholar
Search for Todoroki, Shin-ichi

Publications
Information structure in language acquisition - Publications

Journal article

- Pyykkonen, P., Matthews, D., & Järvič, J. (in press). Three-year-olds are sensitive to semantic prominence during online spoken language comprehension. A visual world study of pronoun resolution. Language and Cognitive Processes. more »
- Narasimhan, B., & Dimroth, C. (2008). Word order and information status in...
Expression Research on Middle Aged Men

<table>
<thead>
<tr>
<th>Album Url</th>
<th><a href="http://test-faces.mpdl.mpg.de/album/escidoc:41047">http://test-faces.mpdl.mpg.de/album/escidoc:41047</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Date Created</td>
<td>Date published: 2009-03-27 at 13:49:11 CET</td>
</tr>
<tr>
<td>Date Created</td>
<td>2009-03-27 at 11:39:59 CET</td>
</tr>
<tr>
<td>Author</td>
<td>Kleinfercher, Friederike Max Planck Digital Library</td>
</tr>
<tr>
<td>Author</td>
<td>Büchner, Kristina Max Planck Digital Library</td>
</tr>
<tr>
<td>Description</td>
<td>This album contains a selection of middle aged men with different facial expressions.</td>
</tr>
</tbody>
</table>

Publication of Research Data
eSciDoc.Faces

Support for discipline-specific images and metadata
Easy to re-use for other disciplines
### eSciDoc.ViRR – Digitized Text Resources

#### Editor

<table>
<thead>
<tr>
<th>Volume</th>
<th>Date Modified</th>
<th>Date Published</th>
<th>Status</th>
<th>Edit Metadata</th>
<th>Edit Table of Contents</th>
<th>Release TOC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zappa, Christian Gottfried</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Schaurosroth, Eberhard Christian Wilhelm von:
Vollständige Sammlung Aller Conclusorum, Schreiben Und anderer übrigen Verhandlungen Des Hochpreußlichen Corporis Evangelicorum. Vol. 1, 1751
Translation of Ideas into Specifications
Creating Initial Artifacts

Preparations
- Narrative explanations (Scenarios)
- Develop awareness about domain standards in parallel
- Get Overview on relevant tools

Details
- Formalized Specifications (Use Cases)
- Functional Mappings
- Investigate on API‘s and formats

While ensuring user involvement
Functional Specification

This specification describes how the yearbook of the MPS will be handled as long as eDoc is still used for the creation of the yearbook, while the rest is handled within PubMan. The idea is, that eDoc should only be used to release the data exported from PubMan for the yearbook of the MPS.

UC_PM_YB_01 Create yearbook container in PubMan

In PubMan a container (folder) per each institute has to be created by the development team manually. In such a container all data which has to be gathered for the MPS yearbook should be collected.

Status / Schedule
Status: In specification
Schedule: R 6.1 (February 2010)

Actors
Development Team

Pre-Conditions
- institute collects/maintains all their metadata and publications in PubMan
- eDoc is just used for the creation of the MPS yearbook

Flow of Events
- the respective institute informs the PubMan support team about the need of the yearbook container in PubMan
- support team helps dev team to create the yearbook container for the MPI XY
- dev team creates the recommended yearbook container manually
  - a container only holds the links to items
  - containers will have metadata in the responsible institutes (OU) via the context

Post-Conditions / Results
- yearbook container is available on PubMan for the respective institute

Future Development
- local administrator of institute is able to create their yearbook container in PubMan self-contained?

UC_PM_YB_02 Add PubMan items to yearbook container

The moderator of the institute chooses the respective items for the MPS yearbook container in PubMan.

Status / Schedule
Status: In specification
Schedule: R 6.1 (February 2010)

Actors
Moderator

Pre-Conditions
- yearbook container for the institute's data has to be created by the dev team beforehand -> see UC_FB_YB_01
Participation

Aim

- Keep development close to researcher needs
- Find local “multipliers” at institute

Target Groups

- MPS institutes (scientists, Library, IT)

Involvement

- Shaping release plans according to needs (Early Adopters)
- Testing, feedback, evaluations (features, GUI)
- First productive users

Organisation/Tools

- Early Adopters/Partners (vital interest, commitment institute)
- Pilot group (fast followers, library know-how)
- Regular workshops, Usability interviews
- CoLaboratory (wiki-based) [http://colab.mpdl.mpg.de/](http://colab.mpdl.mpg.de/)
- Mailinglists
Knowledge gain

Metadata Encoding and Transmission Standard

The METS schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library, expressed using the XML schema language of the World Wide Web Consortium. The standard is maintained in the Network Development and MARC Standards Office of the Library of Congress, and is being developed as an initiative of the Digital Library Federation.

METS is intended to provide a standardized XML format for transmission of complex digital library objects between systems. One METS file corresponds to one digital object (i.e., one digitized work) and provides separate sections for descriptive metadata, administrative metadata, structural metadata, files, and behavior. The structural parts are directly defined by the METS standard, while the other sections incorporate "extension schemas", e.g., MARC/Dublin Core for descriptive metadata or MIX for technical metadata. METS is very powerful for grouping together various digital items into one research object, e.g., to combine scans and TEI transcription of one work.

METS is highly flexible and allows multiple representations of the same digital object. In particular, METS does not restrict the usage of metadata schemas (it only defines a set of supported schemas = "extension schemas") and the structural maps can be organized in multiple ways. Therefore, the standard itself does not provide interoperability. METS profiles may reduce this problem to a certain extend.

METS structure

An example METS xml is available from the Fedora homepage and a METS structure diagram is provided as well.

Header (metsHdr)

Information about the METS document itself, e.g., various time stamps and the institutions and/or individuals (agent) involved in creating the package:

```
< Mets:metsHdr ID="BOOK1" CREATEDATE="2007-05-20T06:32:00" LASTMODDATE="2007-05-22T06:32:00" RECORDSTATUS="A">
  <Mets:agent ROLE="CREATOR" TYPE="ORGANIZATION">
    <Mets:name>Max Planck Institute for History of European Law</Mets:name>
  </Mets:agent>
</Mets:metsHdr>
```
A Process Skeleton to Identify Responsibilities

1. Decision for Scenario
2. Requirements Phase
3. Specification Phase
4. Design Phase
5. Development Phase
6. Staging Phase
7. Production Phase

Steps:
- Agreement on Scope and Approach
- Agreed Input for GUI and SW Development
- Agreed Architecture and GUI Design
- Implemented Scenarios And Documentation
- Proven Maturity of Release
- System in productive Use
Planning tool JIRA – keep track of detailed tasks ...

### Road Map
- **View personal road map**
- **Scope:** next 3 versions | all versions

#### R3.8 (Release Notes)

<table>
<thead>
<tr>
<th>Issue</th>
<th>Status</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS-298</td>
<td>FIXED</td>
<td>Design UC, FM, SM, 09 validate item</td>
</tr>
<tr>
<td>AS-425</td>
<td>FIXED</td>
<td>Check revised design for searching and modify in case needed</td>
</tr>
<tr>
<td>AS-537</td>
<td>FIXED</td>
<td>Test migration of data from currently used build 1.0 beta3 build 237 to pre-release 1.0 beta4 build 306</td>
</tr>
<tr>
<td>AS-488</td>
<td>FIXED</td>
<td>Definition re-definition of mappings eDoc, PubMan</td>
</tr>
<tr>
<td>AS-473</td>
<td>FIXED</td>
<td>Finalize open questions for migration</td>
</tr>
<tr>
<td>AS-467</td>
<td>FIXED</td>
<td>Specification copy, Export</td>
</tr>
<tr>
<td>AS-539</td>
<td>FIXED</td>
<td>Define and implement test cases for search component</td>
</tr>
<tr>
<td>AS-152</td>
<td>FIXED</td>
<td>Revise Architectural Blueprint</td>
</tr>
<tr>
<td>AS-184</td>
<td>WONT FIX</td>
<td>Matrix Actions/Role/Constraints</td>
</tr>
<tr>
<td>AS-522</td>
<td>FIXED</td>
<td>Define requirements for Shikshaleh</td>
</tr>
<tr>
<td>AS-511</td>
<td>FIXED</td>
<td>Input for mapping and open questions for migration</td>
</tr>
<tr>
<td>AS-534</td>
<td>FIXED</td>
<td>Item views should enable multiple lines with scrolling (view All) possibility for longer fields such as Topic, Abstract etc.</td>
</tr>
<tr>
<td>AS-518</td>
<td>FIXED</td>
<td>Run the transformation to item xml + full text</td>
</tr>
</tbody>
</table>

### Preset Filters
- **All**
- **Resolved recently**
- **Outstanding**
- **Added recently**
- **Unscheduled**
- **Updated recently**
- **Assigned to me**
- **Most important**
- **Reported by me**

### Project Summary
- **Closed:** 11 | 2%
- **Not Started:** 57 | 12%
- **Started:** 20 | 4%
- **On Hold:** 15 | 3%
- **Ready:** 58 | 12%
- **Completed:** 320 | 65%
- **Suggested For Rejection:** 3 | 1%
- **Rejected:** 9 | 2%

### Open Issues
- **By Priority**
  - **Blocker:** 4 | 4%
... and overview on projects

<table>
<thead>
<tr>
<th>Key</th>
<th>Components</th>
<th>Fix Version/s</th>
<th>Summary</th>
<th>Status</th>
<th>REQ</th>
<th>SPEC</th>
<th>GUI</th>
<th>DES</th>
<th>CSS</th>
<th>IMP</th>
<th>STA</th>
<th>PROD</th>
<th>Pr</th>
<th>WorkflowActions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP-11</td>
<td>PubMan</td>
<td>R3</td>
<td>Browse &amp; Display</td>
<td>Started</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SetStatus Complete activity Put on hold</td>
</tr>
<tr>
<td>MP-16</td>
<td>PubMan</td>
<td>R3</td>
<td>Context (Collection) Administration</td>
<td>Started</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SetStatus Complete activity Put on hold</td>
</tr>
<tr>
<td>MP-18</td>
<td>PubMan</td>
<td>R3</td>
<td>Controlled Vocabulary Service (Journal names)</td>
<td>Started</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SetStatus Complete activity Put on hold</td>
</tr>
<tr>
<td>MP-8</td>
<td>PubMan</td>
<td>R3</td>
<td>Easy Submission</td>
<td>Started</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SetStatus Complete activity Put on hold</td>
</tr>
<tr>
<td>MP-13</td>
<td>PubMan</td>
<td>R3</td>
<td>Export</td>
<td>Started</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SetStatus Complete activity Put on hold</td>
</tr>
<tr>
<td>MP-19</td>
<td>PubMan</td>
<td>R3</td>
<td>Feeding local web pages - phase 1</td>
<td>Started</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SetStatus Complete activity Put on hold</td>
</tr>
<tr>
<td>MP-14</td>
<td>PubMan</td>
<td>R3</td>
<td>Organisational unit management</td>
<td>Started</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SetStatus Complete activity Put on hold</td>
</tr>
<tr>
<td>MP-32</td>
<td>PubMan</td>
<td>R3</td>
<td>Prototype Citation service</td>
<td>Started</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>SetStatus Complete activity</td>
</tr>
</tbody>
</table>