COMEDI: A New COmponent Metadata EDItor

Gunn Inger Lyse, Paul Meurer, Koenraad De Smedt

CLARINO

University of Bergen / Uni Research Computing

CAC, Soesterberg, October 25, 2014
Outline

Background

Existing editors

COMEDI
Why another metadata editor?

Creating and managing metadata

- is challenging for the inexperienced
- is time-consuming even for the experienced
Why another metadata editor?

Creating and managing metadata

- is challenging for the inexperienced
- is time-consuming even for the experienced

CMDI metadata in particular

- required for CLARIN B centres
- existing editors lack sufficient support for CMDI + user-friendliness
- needs up to date support with CLARIN features (e.g. current license types)
Outline

Background

Existing editors

COMEDI
ARBIL

• Developed at the Max Planck Institute for Psycholinguistics, Nijmegen
• Supports CMDI and IMDI
• For archivists and librarians, not for beginners
• Steep learning curve
ProFormA

- Developed at the University of Tübingen for NaLiDa
- Web-based
- Supports CMDI fully in theory, but..
- Only NaLiDa profiles have been tested
- For other profiles: display errors (cardinality), limitations on controlled vocabulary..
- No validation of content
META-SHARE

- Network of repositories resulting from the META-NET Network of Excellence
- Web-based
- No full CMDI support, only four built-in schemas which can be converted
- Autofill, autocompletion and validation
- Person info, institutional info and project info as reusable objects
- Slightly confusing display and navigation
Outline

Background

Existing editors

COMEDI
COMEDI

- COMEDI=COmponent Metadata EDItor
- Handles any CMDI profile (currently v1.1)
- Web-based
- Upload, edit and download of any record (CMDI XML file; OAI-PMH wrapper)
- Login via eduGAIN, Clarin IdP or OpenIdP
- PID service (URN) (in progress)
- Can easily be deployed as a stand-alone application.
- Available under BSD license
Create new or upload CMDI file

Create a metadata record
Add a resource metadata record, starting from an empty profile

Choose profile: CLARINO corpusProfile – unpublished
... or provide a valid CMDI Profile ID:

Please provide an identifying name for the resource. This name will be used internally.
Identifier: demo for CAC’14

Upload a metadata record
Upload an existing syntactically valid CMDI 1.1 metadata record. The file name (without extension) will be used as identifier.

Choose file: Velg fil Ingen fil valgt

Upload
### Search in metadata records

Freetext search in existing metadata records. All matching records will be listed.

<table>
<thead>
<tr>
<th>Resource name</th>
<th>CMDI Record XML</th>
<th>Owner</th>
</tr>
</thead>
<tbody>
<tr>
<td>Click to view or edit metadata record</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Das große Krötenbuch, mit braunfarbigen Illustrationen. request-test</td>
<td>[Download]</td>
<td>paul</td>
</tr>
<tr>
<td>Koenraads corpus</td>
<td>[Download]</td>
<td>koenraad</td>
</tr>
<tr>
<td>Koenraads corpus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>META-NORD Acquis English Treebank eng-jrc-acquis</td>
<td>[Download]</td>
<td>gunn</td>
</tr>
<tr>
<td>META-NORD Acquis Estonian Treebank est-jrc-acquis-dep</td>
<td>[Download]</td>
<td>gunn</td>
</tr>
</tbody>
</table>
Acess: Read/Write management

COMEDI :: Metadata Editor

Profile ID: clarin.eu:cr1:p_1407745711925; fetch date: 2014-10-14T09:01:17Z
Identifier: request-test (Click to view OAI-PMH record)
Creator: Paul Meurer
Creation Date: 2014-10-10T07:52:39Z
Last Change Date: 2014-10-13T16:04:29Z
Status: published
Access: Read/Write (shared)
Display - view mode

**Contact person** [1-∞]

Role: contactPerson

**Person info** [1]

Surname [nob]: Lyse
Given name [xfew]: Gunn Inger

Lang Validation error: The language code "xfew" does not occur in ISO 639-3.

**Communication info** [1]

Email: [value is missing]
Display - edit mode

**Communication info** [1/1] | ref: - | select existing Communication info | Hide content

Groups information on communication details of a person or an organization

**Email** [1/1-∞] + | [ISOCat]

The email address of a person or an organization

**Url** [0/0-∞] + | [ISOCat]

A URL used as homepage of an entity (e.g. of a person, organization, resource etc.) and/or where an entity document etc. is located
User-friendliness features

- Validation according to ValueScheme specification
- Easy reuse of components from existing metadata records
  - through cloning (editable copies)
  - through sharing (pointer to same component)
- Favourites (like ARBIL)
- Support for referring to resource parts (id/ref) in CMDI, with consistency check
Saving, input and spell checking

- Working copies saved during a working session in the server database,
  - allowing work to be resumed at any time.
  - regardless of whether all user content is valid and complete.
- Unicode allowed for anything typed in
- Spell checking handled by the browser
Support for license information in CLARIN

- Challenge: consistent and correct license info in metadata.
- UiB has classified licences wrt. User Category, License Family, Conditions of Use.
- CLIC (CLARIN Legal Issues Committee) has quality-checked the license table.
- Candidate to be published as external vocabulary in CLAVAS? (supported in CMDI 1.2)
- With CLARINO’s license component, COMEDI adds the correct Conditions of Use automatically.
Support for license information in CLARIN

**Licence** [1/1] | remove content | Hide content

Groups information about a licence governing the use of the resource. restrictionsOfUse in the clarin.eu:cr1:c_1352813745464 component (li

**Licence family** [1/1] | [ISOCat]

CLARIN

**Licence name** [1/1] | [ISOCat]

CLARIN_ACA-NC
Support for license information in CLARIN

**Licence** [1/1] | remove content | Hide content

Groups information about a licence governing the use of the resource. restrictionsOfUse in the clarin.eu:cr1:c_1352813745464 component (li

---

**Licence family** [1/1] | [ISOCat]

<table>
<thead>
<tr>
<th>CLARIN</th>
</tr>
</thead>
</table>

---

**Licence name** [1/1] | [ISOCat]

<table>
<thead>
<tr>
<th>CLARIN_ACA-NC</th>
</tr>
</thead>
</table>

---

**Conditions of use** [4/1-∞] | [ISOCat]

Corresponds to the META-SHARE "restrictionsOfUse". Changed the n restrictions* (e.g. "commercial use")

- BY
- ID
- NC
- NORED
COMEDI: first user test

- Low threshold for first use
- Top-level component menu is useful
- Autosave function when switching from Edit to View is desired
- More user feedback needed
COMEDI demo

Demo session this afternoon!

http://clarino.uib.no/comedi/