Standards and the Standards Committee

Katrien Depuydt
Standards in Elexis

Standards for the lexicographical data in Elexis
- Interoperability
- Open
- Efficient
→ Part of best practices
Interoperability in Elexis

Common protocols for input data

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731015.
Chosen Standards

• TEI Lex-0 (Presentation by T. Tasovac)
• Ontolex-Lemon (Presentation by John McCrae)
Semantic Interoperability

TEI Lex-0 → DATAMODEL → Ontolex-Lemon

Djed Pillar ~ BACKBONE of Osiris ~ stability

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731015.
ELEXIS data model

- a **minimal common data model** capturing the main concepts of lexicographic resources
- takes into account international standards.
- can be expressed by both interoperability formats (Ontolex-Lemon and TEI Lex-0)

The main goal is to ensure **semantic interoperability** between lexicographic resources to enable **integration** in the ELEXIS infrastructure.
Standards Committee in Elexis

Grant Agreement

The aim of SC will be to ensure the realisation of the project Objective 2 by working on common standards and solutions for the development of lexicographic resources.

Objective 2:

establish common standards and solutions for the development of lexicographic resources
<table>
<thead>
<tr>
<th>Target outcome</th>
<th>Indicator</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Publication of standardisation documentation for lexicographic resources</td>
<td>Developed standard included in ISO or OASIS (end of project)</td>
<td>&gt;= 1 new standard</td>
</tr>
<tr>
<td>Use of Data Seal of Compliance</td>
<td>Number of lexicographic resources with the Seal (end of project)</td>
<td>&lt; 100 (not satisfactory)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt; 100 (satisfactory)</td>
</tr>
<tr>
<td>Use of conversion tools to harmonise the different data formats of lexicographic resources (Virtual Access)</td>
<td>Number of resources using API pipeline in LEX1 and LEX3 platforms</td>
<td>&lt; 500 (not satisfactory)</td>
</tr>
<tr>
<td></td>
<td>Number of languages (X)</td>
<td>&gt; 500 (satisfactory)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X &lt; 10 (not satisfactory)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 &lt; X &lt; 20 (satisfactory)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>X &gt; 20 (very satisfactory)</td>
</tr>
<tr>
<td>Scientific output on standardisation in lexicography</td>
<td>Number of scientific journal articles</td>
<td>&lt; 3 (not satisfactory)</td>
</tr>
<tr>
<td></td>
<td>Number of scientific conference papers</td>
<td>&gt; 3 (satisfactory)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt; 15 (not satisfactory)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&gt; 15 (satisfactory)</td>
</tr>
</tbody>
</table>
Standards Committee meeting 18-2-2019

• Presentations on
  • Existing standards and ELEXIS: based on FCSH-UNL Report on ISO Standards for ELEXIS TEI Lex-0 and ELEXIS
  • Ontolex-Lemon and ELEXIS
  • ELEXIS data model
• Discussion on the essential role of the Elexis data model and the possibility to make the model a standard in OASIS, with both Ontolex Lemon and TEI Lex-0 as serialisations
• Plans for joined Ontolex-Lemon TEI Lex-0 meeting
Members of the Standard Committee

Tomaž Erjavec, JSI – standard: TEI, also member of CLARIN Standards Committee
Carole Tiberius, INT – standard: TEI, LMF, Ontolex-Lemon
Katrien Depuydt, INT – standard: TEI, Ontolex-Lemon
Toma Tasovac, BCDH – standard: TEI
Fahad Khan, CNR-ILC – standard: LMF, also member of CLARIN Standards Committee
Monica Monachini, CNR-ILC – standard: LMF
John McCrae, NUIG – standard: Ontolex-Lemon
Thierry Declerck, OEAW – standard: Ontolex-Lemon
Karlheinz Mörth, OEAW, also member of CLARIN Standards Committee
David Filip, LC – standardisation organisation: OASIS
Rute Costa, FCSH-UNL – standardisation organisation: ISO
Christophe Roche, FCSH-UNL
Simon Krek, JSI – standard: TEI
Yifat Ben Moshe, KD
Miloš Jakubíček, LC

This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 731015.