Multilingual Knowledge Systems as Linguistic Linked Open Data for European Language Grid

SALLD-1 (Workshop on Sentiment Analysis & Linguistic Linked Data) – 01 September 2021
Agenda

Coreon MKS
MKS2LLOD
Impact
Coreon: Multilingual Knowledge System

- Combines knowledge graphs (aka *concept systems*) with multilingual terminology
- MKS = semantic bridge between different systems, organizations, cultures
- MKS = systematic terminology management = multilingual knowledge graph
- Software-as-a-service solution, deployed for large scale corporates and institutions
**Multilingual Knowledge System (MKS):**
Terminology and Knowledge Graph in One

MKS: Delivers what a feature-rich stand-alone terminology management tool delivers.

MKS: Models what taxonomy / thesaurus / ontology editors model.

… thus Coreon combines terminology with the advantages of a knowledge structure …

Florian Deubzer, SDI Munich
Great for Maintainers: Terminologists, Ontologists ...
- Work in **context**, see the gaps
- Work **systematically**, avoid doubletes, keep control
- Capture both, the **lexical** and the **encyclopedic** aspect

Great for the Audience in Organisations and Businesses
- Share your **concept system in all languages**
- Simply with a browser, in a great UI
Building Blocks (before MKS2LLOD)
MKS2LLOD Project

- Instead of using proprietary APIs, achieve high efficiency by complementing MKS with standards known from LLOD (Linguistic Linked Open Data).
- Enable companies and institutions to make their MKS (= semantic interoperability assets) discoverable and available through ELG.
- Develop and complement Coreon with a SPARQL interface, API key based authentication and granular access rights management.
- ELG to be a gateway to the SPARQL endpoint = become a platform for the multilingual semantic web.
The Dilemma Up to Today

Knowledge Graphs: SPARQL & RDF
- Not appropriate neither designed for terminology

Termbases: Proprietary APIs or TBX
- Not for the Semantic Web
- Tedious exporting / re-modeling, deploy and import RDF triple store

Lost opportunity to leverage terminology data for Semantic Web solutions
Building Blocks (enriched with MKS2LLOD)

- **Authentication Service**
  - Ruby on Rails
- **Coreon Service**
  - Ruby on Rails
- **Coreon User Store**
  - MongoDB
- **Coreon Data**
  - MongoDB
- **Coreon Index**
  - Apache Solr
- **Coreon Search**
  - ESTeam ALiSe
- **Coreon Triple Store**
  - Apache Jena Fuseki
- **RESTful Web**
- **Browser: Coreon UI**
  - Javascript
- **Custom App**
  - .Net, Java, ...
- **Semantic Web Solution**
  - SPARQL consumer ...
- **SPARQL endpoint**
  - RDF
- **Other IDP**
  - <Directory Svc
- **SAMLI2**

Coreon Cloud
Required Activities to Get there

1. Remodel Coreon data structures into an RDF vocabulary
2. Build a serializer to export in various syntax variants
   1. Select an RDF Triple Store solution
   2. Build a mirroring mechanism into this RDF store
   3. Make the whole module part of Coreon’s cloud deployment

1. For easy access: Complement Coreon authentication module with an API key based mechanism
Dilemma Resolved

Query terminology with the syntax known from the Semantic Web

Real-time
No need to export, mirror into, nor deploy a separate triple store

Findable
Promoted and testable on the ELG hub

Shared
Best practices and learnings feed into an ISO Technical Recommendation

Make Terminology Resources a Component of the Semantic Web
Agenda

Coreon MKS

MKS2LLLOD

Impact
Impact 1 / 2: Technology Showcase and Best Practices

- SPARQL endpoint to real time
  Coreon MKS
- Integrated in ELG
- Standards best practice: modelling terminology in RDF
Impact 2 / 2: Building Bridges in Businesses and Organisations

Public Institutions
• enable cross-border semantic interoperability

Corporates
• push concepts and terms into business ecosystems

Achieve “FAIR”ness
• Findable, Accessible, Interoperable, Reusable
The MKS2LLOD project was supported by the European Language Grid project through its open call for pilot projects. The European Language Grid project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under Grant Agreement no. 825627 (ELG).