Linked Data and Language Technologies: The LIDER project

A. Gómez-Pérez (UPM)

asun@fi.upm.es

Project Coordinator

CSA
Budget: 1.482.000€
Starting date: 1. Nov. 2013
Duration: 2 Years
163 PM
The paradigm and business model have changed.
14.05.14

Asun Gómez-Pérez

Technical activities

1. Roadmap on 3LD for Content Analytics
2. Guidelines for 3LD
3. 3LD Reference Architecture

Industry use cases

Community building networking

LD4LT
BP-MLOD W3C-CG
OntoLex W3C-CG

.- Surveys
.- Requirements
Asun Gómez-Pérez

Architecture

Multimedia and Multilingual Content

Metadata Generation

Content Analytics

LOD-aware NLP services

Language resources as LD

Linguistic LOD generation (Metadata and Content)

Language Resources (Lexicon, corpora, ...) some of them are FOI other are private
A lifecycle for LRs Metadata

Metadata conversion

Publication

PROVENANCE AND LICENSE

Indexing

Enhanced navigation and discovery of LRs

Other sources

Linguistic LD Observatory

XML

META-SHARE

META-SHARE

RDF

META-SHARE

LinguisNc

LD Observatory

Observatory
Vocabularies for describing Linguistic Data

General technical metadata
(Resource names, maintainers, languages, formats, data links)

License metadata
(openness, editability, attribution)

Domain specific metadata
(repository & resource specific vocabularies and ontologies)

Core
- DCAT,
- VoID, ..
- ODLR,
- Prov-O, ..
- Metashare,
- Datahub model /
  Lemon, NIF,
  MARL, OliA, ...

Conclusions

What META-SHARE can do for LIDER

• Help LIDER not to “reinvent the wheel”
• Provide a critical mass of linguistic metadata/data in RDF
• First step towards a new generation of LD-aware NLP methods and tools

What LIDER can do for META-SHARE

• Promote MS core metadata in the LD4LT @W3C
• Help exposing metadata and linking LRs as LD, which is an unstoppable trend
• Help to curate the MS metadata where exposing metadata in RDF
• Expose provenance and license information in a machine readable way
• Enhancement of current MS infrastructure with an "extra layer" to enable easier ways of resource discovery
• Bi-directional harvesting
Linked Data is to be accessed and used by machines.
Web channels

Join the community
www.w3c.org/community/ld4lt

twitter.com/multilingweb
Hashtag: #LiderEU