



## LIDER – Linked Data as an enabler of cross-media and multilingual content analytics for enterprises across Europe

### Challenge

The explosive growth of content in volume, velocity and variety on the Web demands new approaches to content analytics. This is the prerequisite for addressing issues in large-scale analysis and interpretation of heterogeneous data sets, originating in different media, human languages or organisations.

Recent advances in and uptake of Linked Data and Semantic Web technology has improved the interoperability needed for different tools to easily conduct “big data” analytics over the torrents of structured content and social media generated daily by organisations and their users. However, a key missing element is the understanding and interpretation of language in content, both as unstructured textual content and linguistic content present in the context of different media streams. The representation of language- and media-specific linguistic information on a semantic level is needed – accurate analytics operating across the increasing variety of media and human languages used nowadays on the Web.

### Objective and Innovation

LIDER will build a pathway for the establishment of a new Linked Open Data (LOD) based ecosystem of free, interlinked, and semantically interoperable resources from the realm of both language (“Linguistic Linked Data” representations of corpora, dictionaries, lexical and syntactic metadata, etc.) and media (image, video, etc. metadata). LIDER will study how these resources can serve as an enabler technology for enterprise content analytics on the Multilingual Web, including multilingual content delivered in

multiple media.

LIDER will help to build the community around the Linguistic Linked Licensed Data (3LD) cloud that includes open linguist resources and licensed linguistic resources in the linked data format.

### Target groups

A wide range of communities will benefit from LIDER: consumers and providers of multilingual and multimedia content, multinational enterprises, European public bodies and SMEs, Language Service Providers and the language technology research and industry communities at large. LIDER will actively engage with these communities to elicit the specific business cases and use cases that will drive the outputs of the project.

### Expected result

LIDER will create:

- A strong community around the topic of LOD-based multimedia and multilingual content analytics.
- A set of guidelines and best practices for building and exploiting LOD-based resources in multimedia and multilingual content analytics and for developing NLP services on top of Linguistic Linked Data.
- A reference architecture for Linguistic Linked Data built on top of existing and future platforms and freely available resources.
- A long-term roadmap for the use of Linked Data for multilingual and multimedia content analytics in enterprises.



## LIDER – Linked Data as an enabler of cross-media and multilingual content analytics for enterprises across Europe

### Impact

LIDER will enable the afore-mentioned stakeholders to develop a shared understanding about the representation of language- and media-specific information on a semantic level. The long-term roadmap will receive input from a significant number of companies and research organizations. In this way, LIDER firstly will lay the groundwork for reducing the costs of adapting an existing analytics solution to multiple languages and across media boundaries. Secondly, the LIDER roadmap will positively influence the involved centres, by contributing to scientific excellence and coordination while assuring the industrial relevance of their research activities.

### Coordinator

Contact Person: Prof. Dr. Asunción Gómez-Pérez

- Tel:+34 913367417
- E-mail: [asun@fi.upm.es](mailto:asun@fi.upm.es)
- Organisation: Universidad Politécnica de Madrid (UPM)

### Participants

- Universidad Politécnica de Madrid (UPM)
- Trinity College Dublin (TCD)
- German Research Center for Artificial Intelligence (DFKI)
- INSIGHT National Center for Data Analytics, National University of Ireland, Galway
- Institute for Applied Informatics (InfAI), University of Leipzig
- University of Bielefeld
- Sapienza University of Rome
- W3C / Geie ERCIM

### Project setup

- Duration: 24 months
- Start date: 1 November 2013
- End Date: 31 October 2015
- Project Officer: Susan Fraser
- Project 610782, CSA

## How to contribute and how to be informed about LIDER

- Project website <http://www.lider-project.eu/>
- Community portal <http://www.multilingualweb.eu/>
- Upcoming W3C Community Group on “Linked Data for Language Technology”
- Mailing list [lider-community@listas.fi.upm.es](mailto:lider-community@listas.fi.upm.es) for gathering anonymous feedback and for getting information. Use <http://www.lider-project.eu/get-involved> to subscribe
- Social media presence via hastagh #lider-project e.g. on twitter <https://twitter.com/multilingweb>