



LIDER: FP7 – 610782

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

Deliverable number **D4.4.2**

Deliverable title **Updated Project Fact Sheet,
Phase I**

Main Authors **Felix Sasaki**

Grant Agreement number	610782
Project ref. no	FP7-610782
Project acronym	LIDER
Project full name	Linked Data as an enabler of cross-media and multilingual content analytics for enterprises across Europe
Starting date (dur.)	1/11/2013 (24 months)
Ending date	31/10/2015
Project website	http://www.lider-project.eu/

Coordinator	Asunción Gómez-Pérez
Address	Campus de Montegancedo sn. 28660 Boadilla del Monte, Madrid, Spain
Reply to	asun@fi.upm.es
Phone	+34-91-336-7417
Fax	+34-91-3524819

Document Identifier	D4.4.2
Class Deliverable	LIDER EU-ICT-2013-610782
Version	1.0
Document due date	31 October 2014
Submitted	5 November
Responsible	W3C/ERCIM
Reply to	fsasaki@w3.org
Document status	Final
Nature	O(Other)
Dissemination level	PU(Public)
WP/Task responsible(s)	Felix Sasaki, DFKI / W3C Fellow
Contributors	-
Distribution List	Consortium Partners
Reviewers	Reviewed by the project consortium
Document Location	http://lider-project.eu/?q=doc/deliverables

Executive Summary

This document, the updated Project Fact Sheet – Phase I, summarizes basic information about the LIDER project and highlights the outcome of the first 12 months of the project. Another updated version will be published in month 24.

The main pages of this deliverable provide the textual content of the Project Fact Sheet as a text document. The final pages provide the Project Fact Sheet visualized as a flyer for printing.

Document Information

IST Project Number	FP7-610782	Acronym	LIDER
Full Title	Linked Data as an enabler of cross-media and multilingual content analytics for enterprises across Europe		
Project URL	http://www.lider-project.eu/		
Document URL	http://lider-project.eu/?q=doc/deliverables		
EU Project Officer	Susan Fraser		

Deliverable	Number	D4.4.2	Title	Update Project Fact Sheet, Phase I
Workpackage	Number	4	Title	Community building and dissemination

Date of Delivery	Contractual	31 October 2014	Actual	5 November
Status	version 1.1		Final ■	
Nature	prototype <input type="checkbox"/> report <input type="checkbox"/> dissemination ■			
Dissemination level	public ■ consortium <input type="checkbox"/>			

Authors (Partner)	Felix Sasaki, DFKI / W3C Fellow			
Responsible Author	Name	Felix Sasaki	E-mail	fsasaki@w3.org
	Partner	DFKI / W3C Fellow	Phone	+49-30-23895-1807

Abstract (for dissemination)	This document, the updated Project Fact Sheet – Phase I, summarizes basic information about the LIDER project and highlights the outcome of the first 12 months of the project. Another updated version will be published in month 24. The main pages of this deliverable provide the textual content of the Project Fact Sheet as a text document. The final pages provide the Project Fact Sheet visualized as a flyer for printing.
Keywords	LIDER, project fact sheet

Version	Modification(s)	Date	Author(s)
01	Initial version	25/10/14	Felix Sasaki, DFKI / W3C Fellow
02	Final version	04/11/14	Felix Sasaki, implementing comments from TCD

Project Consortium Information

Participants		Contact
Universidad Politécnica de Madrid		Asunción Gómez-Pérez Email: asun@fi.upm.es
The Provost, Fellows, Foundation Scholars & The Other Members of Board of The College of the Holy & Undivided Trinity of Queen Elizabeth near Dublin (Trinity College Dubl, Ireland)		David Lewis Email: dave.lewis@cs.tcd.ie
Deutsches Forschungszentrum für Künstliche Intelligenz GmbH (DFKI, Germany)		Felix Sasaki Email: felix.sasaki@dfki.de
National University of Ireland, Galway (NUI Galway, Ireland)		Paul Buitelaar Email: paul.buitelaar@deri.org
Institut für Angewandte Informatik EV (INFAI, Germany)		Sebastian Hellmann Email: hellmann@informatik.uni-leipzig.de
Universität Bielefeld (UNIBI, Germany)		Philipp Cimiano Email: cimiano@cit-ec.uni-bielefeld.de
Universita degli Studi di Roma La Sapienza (UNIVERSITA DEGLI STU, Italy)		Roberto Navigli Email: navigli@di.uniroma1.it
GEIE ERCIM (ERCIM, France)		Felix Sasaki Email: fsasaki@w3.org

Table of Contents

1	BACKGROUND	7
2	MAIN ACHIEVEMENTS IN YEAR 1 OF LIDER	7
	REACHING OUT TO COMMUNITIES, BUILDING BUSINESS USE CASES	7
	GUIDELINES AND BEST PRACTICES.....	7
	REFERENCE ARCHITECTURE AND ROADMAPPING	7
3	NEXT STEPS.....	8
4	COORDINATOR.....	8
5	PARTICIPANTS.....	8
6	PROJECT SETUP	8
7	HOW TO CONTRIBUTE AND HOW TO BE INFORMED ABOUT LIDER	8

1 Background

The aim of LIDER is to address the challenge of the explosive growth of content in volume, velocity and variety on the Web, which demands new approaches to content analytics. A key 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.

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.

2 Main achievements in Year 1 of LIDER

Reaching out to communities, building business use cases

LIDER has held four roadmapping workshops to gather input for business use cases and requirements for linguistic linked data. The workshops targeted four groups: the data community, the Web community at large, the localisation community and the analytics community. The feedback was gathered via a dedicated W3C community group LD4LT (“Linguistic Linked Data for Language Technology”), has been documented in workshop reports and has fed into the LIDER roadmap. In addition, LIDER has been present at a large number of research and industry events, held tutorials and gave presentations about linguistic linked data and content analytics.

Guidelines and best practices

LIDER has used a dedicated W3C community group BPMLOD (“Best Practices for Multilingual Linked Open Data”) to gather best practices for creating multilingual linked data sources. Conversions for existing language resources into linked data have been developed. This concerns first metadata about language resources. LIDER worked closely with the META-NET and the CLARIN community on the conversions, to assure broad acceptance and distribution of existing metadata schemes. Second, LIDER described best practices for various types of resources, like bilingual or multilingual dictionaries, and started conversions for other resources like terminology (TBX) data or corpus data.

Reference architecture and roadmapping

LIDER has created a first draft of a linguistic linked data reference architecture for content analytics. The architecture defines a general model for building linguistic linked data aware services and several patterns for building content analytics applications. It also provides an overview tools and initiatives relevant for the architecture.

LIDER also has created a first version of a roadmap for linguistic linked data for content analytics. The roadmap interprets input from various sources: LIDER activities like surveys or roadmapping workshops, various reports, funding agencies and various research communities. The roadmap is formulated around three key application areas:

global customer engagement, public sector / civil society, and linguistic linked data life cycle and linguistic linked data value chain. For each of these areas, the roadmap provides several use cases including time lines, predictions and relevant actors.

3 Next steps

LIDER will validate its year one outcomes by discussing them with the community at large. This concerns conversion of linguistic resources and its metadata, the linguistic linked data reference architecture for content analytics, and the roadmap for linguistic linked data for content analytics. In addition, LIDER will reach out to new groups that have not been targeted in year one, especially in the area of public sector information.

4 Coordinator

Contact Person: Prof. Dr. Asunción Gómez-Pérez

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

5 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

6 Project setup

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

7 How to contribute and how to be informed about LIDER

- Project website <http://www.lider-project.eu/>
- Community portal <http://www.multilingualweb.eu/>
- W3C Community Group on “Linked Data for Language Technology”
<https://www.w3.org/community/ld4lt/>
- Mailing list 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 hashtag #LiderEU e.g. on twitter
<https://twitter.com/multilingweb>



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

Background

The aim of LIDER is to address the challenge of the explosive growth of content in volume, velocity and variety on the Web, which demands new approaches to content analytics. A key 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. 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: Main achievements in Year 1

➤ **Reaching out to communities, building business use cases.** LIDER has held four roadmapping workshops to gather input for business use cases and requirements for linguistic linked data. The workshops targeted four groups: the data community, the Web community at large, the localisation community and the analytics community. The feedback was gathered via a dedicated W3C community group LD4LT (“Linguistic Linked Data for Language Technology”), has been documented in workshop reports and has fed into the LIDER roadmap. In addition, LIDER has been present at a large number of research and industry events, held

tutorials and gave presentations about linguistic linked data and content analytics.

➤ **Guidelines and best practices.** LIDER has used a dedicated W3C community group BPMLOD (“Best Practices for Multilingual Linked Open Data”) to gather best practices for creating multilingual linked data sources. Conversions for existing language resources into linked data have been developed. This concerns first metadata about language resources. LIDER worked closely with the META-NET and the CLARIN community on the conversions, to assure broad acceptance and distribution of existing metadata schemes. Second, LIDER described best practices for various types of resources, like bilingual or multilingual dictionaries, and started conversions for other resources like terminology (TBX) data or corpus data.

➤ **Reference architecture and roadmapping.** LIDER has created a first draft of a linguistic linked data reference architecture for content analytics. The architecture defines a general model for building linguistic linked data aware services and several patterns for building content analytics applications. It also provides an overview tools and initiatives relevant for the architecture. LIDER also has created a first version of a roadmap for linguistic linked data for content analytics. The roadmap interprets input from various sources: LIDER activities like surveys or roadmapping workshops, various reports, funding agencies and various research communities. The roadmap is formulated around three key application areas:



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

global customer engagement, public sector / civil society, and linguistic linked data life cycle and linguistic linked data value chain. For each of these areas, the roadmap provides several use cases including time lines, predictions and relevant actors.

Next steps

LIDER will validate its year one outcomes by discussing them with the community at large. This concerns conversion of linguistic resources and its metadata, the linguistic linked data reference architecture for content analytics, and the roadmap for linguistic linked data for content analytics. In addition, LIDER will reach out to new groups that have not been targeted in year one, especially in the area of public sector information.

Coordinator

Contact Person: Prof. Dr. Asunción Gómez-Pérez

- Tel:+34 913367417
- E-mail: 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 for gathering anonymous feedback and for getting information. Use <http://www.lider-project.eu/get-involved> to subscribe
- Social media presence via hashtag #LiderEU e.g. on twitter <https://twitter.com/multilingweb>