



LIDER: FP7 – 610782

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

Deliverable number	D4.3.2
Deliverable title	Report on LIDER Community and community portal – Phase I
Main Authors	David Lewis, Asunción Gómez-Pérez , Kevin Koidl, 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	+34-91-3367417
Fax	+34-91-3524819

Document Identifier	D4.3.2
Class Deliverable	LIDER EU-ICT-2013-610782
Version	2
Document due date	30/10/2014
Submitted	8/11/2014
Responsible	David Lewis, TCD
Reply to	Dave.lewis@scss.tcd.ie
Document status	Final
Nature	P(Prototype)
Dissemination level	PU(Public)
WP/Task responsible(s)	WP4/T4.2
Contributors	David Lewis (TCD), Asunción Gómez-Pérez (UPM), Kevin Koidl (TCD), Felix Sasaki (DFKI/ERCIM)
Distribution List	Consortium Partners
Reviewers	INFIA
Document Location	http://lider-project.eu/?q=doc/deliverables

Executive Summary

This document provides an overview of the first year operation of the LIDER community portal. Community engagement has primarily been conducted through a W3C Community Group called Linked Data for Language Technology, though the broader output of the project is also published through a LIDER project web site. Other existing community portals have also been used for engagement on LIDER activities. This document describes the operation of these portals in the first phase of LIDER.

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	D4.3.2	Title	Report on LIDER Community and community portal –Phase I
Workpackage	Number	WP4	Title	Community building and dissemination

Date of Delivery	Contractual	31/1/14	Actual	8/11/2014
Status	version 2		final ■	
Nature	prototype <input type="checkbox"/> report ■ dissemination <input type="checkbox"/>			
Dissemination level	public ■ consortium <input type="checkbox"/>			

Authors (Partner)	David Lewis (TCD), Asunción Gómez-Pérez (UPM), Kevin Koidl (TCD), Felix Sasaki (DFKI/ERCIM)			
Responsible Author	Name	David Lewis	E-mail	Dave.lewis@scss.tcd.ie
	Partner	TCD	Phone	+35318968428

Abstract (for dissemination)	This document provides an overview of the first year operation of the LIDER community portal. Community engagement has primarily been conducted through a W3C Community Group called Linked Data for Language Technology, though the broader output of the project is also published through a LIDER project web site. This document describes the operation of both portals in the first phase of LIDER.
Keywords	Linked Data for Language Technologies, Web Portal, LIDER web site

Version	Modification(s)	Date	Author(s)
01	First draft	3/11/2014	D.Lewis TCD
02	Revised from review	8/11/2014	D.Lewis 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: fsasaki@w3.org
National University of Ireland, Galway (NUI Galway, Ireland)		Paul Buitelaar Email: paul.buitelaar@deri.org
Institut für Angewandte Informatik EV (INFIA, 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	OBJECTIVES	8
2	ENGAGEMENT REPORT: LD4LT COMMUNITY	9
3	ENGAGEMENT REPORT: BPMLOD COMMUNITY	10
4	ENGAGEMENT REPORT: ONTOLEX COMMUNITY	11
5	ENGAGEMENT REPORT: ITS INTEREST GROUP	12
6	ENGAGEMENT REPORT: LIDER PROJECT PORTAL	13

1 Objectives

The Objective of the LIDER Community Portal is to provide an online infrastructure to support the external community that LIDER assembles to engage on requirements and use cases for developing guidelines, best practice and the research roadmap for linguistic linked data.

The LIDER community is also referred to as the Industrial Board, to indicate the intention that it should be largely made up of industrial representatives, and in particular SMEs.

The engagement with and growth of the LIDER Community is conducted through the following channels:

- Linked Data for Language Technology (LD4LT) W3C Community Group at <http://www.w3.org/community/ld4lt/>. This is an open community with a web portal and email lists hosted by the W3C and is the primary means of community engagement for LIDER. It focuses on gathering of requirements and use cases, collecting feedback on technical architecture and research road-mapping activities and consensus building on specific technical specifications.
- In addition, engagement with the LIDER community is conducted through three other communities hosted by the W3C:
 - BPMLD (“Best Practices for Multilingual Linked Open Data”) at, <http://www.w3.org/community/bpmlod/> for topics related to WP2 “Guidelines and best practices for industry”,
 - OntoLex at <http://www.w3.org/community/ontolex/> for developing a linked data based ontology lexicon model
 - ITS Interest Group at <http://www.w3.org/International/its/ig/> for discussing the relation between linked data and localization workflow metadata (ITS 2.0).
- The LIDER project portal hosted by UPM at www.lider-project.eu. This integrates the project’s public announcement mailing list (offered to those not willing to join the LD4LT community portal) and social media feeds together with public GitHub and data repository links.
- The Multilingual Web portal operated by the W3C for making announcements and publishing material from the Multilingual Web conference series. This consists of the www.multilinguaweb.eu web site together with the associated email list (549 subscribers) and social media feeds. Major announcements are also distributed via the W3C home page, thus offering extremely wide visibility

This document provides an overview and usage figures for the first year operation of the LD4LT and the LIDER web portals.

A fully summary of engagement with the target community, including via the face to face engagement through organized events is available as LIDER deliverable D4.2.1. More details of LIDER support for the different online communities is also captured in D4.2.1.

2 Engagement Report: LD4LT Community

The LIDER Community Portal is implemented technically and organizationally as a W3C Community Group. These groups are supported by the W3C to enable community discussion about technical issues that could potentially contribute to standardization on the Web¹. Members require a freely available W3C member account. Community members are free to structure their groups as they see fit amongst themselves, selecting group chairs, discussion/meeting frequency, setting objectives and work timetables and are not subject to control or supervision by the W3C.

The Linked Data for Language Technology (LD4LT) Community Group was set up with the following remit:

“This group aims to consult with current and potential users of linguistic data to assemble user cases and requirements for Language Technology Applications that use Linked Data. The results will be used to guide future interoperability, research and development activities spanning the language technology and linked data domains, including via the OntoLex and BP-MLOD community groups.

Potential users are companies and public bodies involved in content management, the language services and localisation industry and other applications of content analytics techniques used in search, recommender systems, sentiment analysis and terminology management.

The group will engage with users through surveys and international road-mapping events organised by an EU-funded R&D consortium called LIDER. Contributions can be provided openly through engagement with the Community Group or anonymously if preferred. The anonymised outcomes of this exercise, in terms of use case and requirements priorities, technology gaps and interoperability roadblocks will be made publicly available via this community group by May 2014.”

The community currently has 79 members². It is co-chaired both by LIDER project members and external experts. Community engagement is via:

- An email list for meeting announcement and technical discussions³
- Three face to face roadmapping workshops, the proceedings of which are recorded on the LD4LT wiki⁴
- Summary of requirements and use cases gathered to date on the wiki⁵
- Specific technical consensus building material. To date this has focused on linked data meta-data for language resources in collaboration with the META-SHARE project⁶

Below key statistics from the LD4LT wiki for the first year of operation:

¹ <http://www.w3.org/community/about/#cg>

² <http://www.w3.org/community/ld4lt/participants>

³ <http://lists.w3.org/Archives/Public/public-ld4lt/>

⁴ https://www.w3.org/community/ld4lt/wiki/Lider_roadmapping_activities

⁵ https://www.w3.org/community/ld4lt/wiki/Main_Page#Use_Cases

⁶ https://www.w3.org/community/ld4lt/wiki/Meta-Share_OWL_metamodel

Page statistics	
Content pages	7
Pages (All pages in the wiki, including talk pages, redirects, etc.)	81
Uploaded files	60
Edit statistics	
Page edits since Linked Data for Language Technology Community Group was set up	754
Average edits per page	9.31
User statistics	
Registered users	18
Active users (list of members) (Users who have performed an action in the last 30 days)	9
View statistics	
Views total (Views to non-existing pages and special pages are not included)	77,118
Views per edit	102.28
Most viewed pages	
LD4LT Group Kick-Off and Roadmap Meeting	50,350
LD4LT Group Madrid May 2014 Meeting	7,233
Main Page	6,780
Seed Use Cases	4,727
LD4LT Group Kick-Off and Roadmap Meeting Agenda	1,252
Meta-Share OWL metamodel	643
Teleconference logistics	573
XML Prague 2014 report	568
LD4LT Roadmapping Workshop Dublin June14	528
Linguistic Linked Data for Content Analytics: a Roadmap	371

3 Engagement Report: BPMLOD Community

BPMLOD describes best practices for multilingual linked data, including conversions of existing data into linked data. These best practices include work conducted in WP2 of LIDER together with other community input. Its outputs are also fed to the Best Practices group of the W3C Data Activity⁷.

Below are key statistics from the BPMLOD Community wiki. As this group existed before the start of the project they cover a period longer than the first phase of LIDER:

Page statistics

⁷ http://www.w3.org/2013/dwbp/wiki/Main_Page

Content pages	15
Pages (All pages in the wiki, including talk pages, redirects, etc.)	35
Uploaded files	3
Edit statistics	
Page edits since Best Practices for Multilingual Linked Open Data Community Group was set up	431
Average edits per page	12.31
User statistics	
Registered users	17
Active users (list of members) (Users who have performed an action in the last 30 days)	4
View statistics	
Views total (Views to non-existing pages and special pages are not included)	5,966
Views per edit	13.84
Most viewed pages	
Main Page	1,354
Best practises - previous notes	587
Use cases definition	572
Topic classification	494
Meetings of the community group	379
Group charter	311
Case studies	289
Guidelines for LD generation of Language resources - previous notes	190
Teleconference, 17/01/2014	119
Kick-off teleconference, 2013.20.06, 9-10 am CET	113

4 Engagement Report: OntoLex Community

The OntoLex community is building ontology lexicon models.

Below are key statistics from the OntoLex community wiki. As this group existed before the start of the project they cover a period longer than the first phase of LIDER:

Page statistics	
Content pages	21
Pages (All pages in the wiki, including talk pages, redirects, etc.)	130
Uploaded files	9
Edit statistics	

Page edits since Ontology-Lexica Community Group was set up	1,510
Average edits per page	11.62
User statistics	
Registered users	20
Active users (list of members) (Users who have performed an action in the last 30 days)	4
View statistics	
Views total (Views to non-existing pages and special pages are not included)	30,637
Views per edit	20.29
Most viewed pages	
Main Page	3,981
Specification of Use Cases	2,795
Final Model Specification	1,095
Specification of Requirements	1,014
Specification of Requirements/Linked Data	1,007
Specification of Requirements/Lexicon-Ontology-Mapping	964
Goals and Scope of Ontology-Lexica Community Group	828
Specification of Requirements/Properties-and-Relations-of-Entries	800
Eurosentiment	612
Core Model Poll	599

5 Engagement Report: ITS Interest Group

The ITS Interest Group is a chartered W3C Interest Group under the Internationalization Activity⁸ at the W3C. Its mission is to foster a community of users of the Internationalization Tag Set (ITS) Version 1.0⁹ and ITS 2.0¹⁰, by promoting its adoption within and outside of W3C, and gathering information on its further development. Of relevant to LIDER is its work on the ITS mapping to RDF¹¹ and Open Data Management of Public Automated Translation Services¹².

Below are key statistics from the ITS IG wiki. As this group existed before the start of the project they cover a period longer than the first phase of LIDER:

Page statistics	
Content pages	17

⁸ <http://www.w3.org/International/>

⁹ <http://www.w3.org/TR/its/>

¹⁰ <http://www.w3.org/TR/its20/>

¹¹ http://www.w3.org/International/its/wiki/ITS-RDF_mapping

¹²

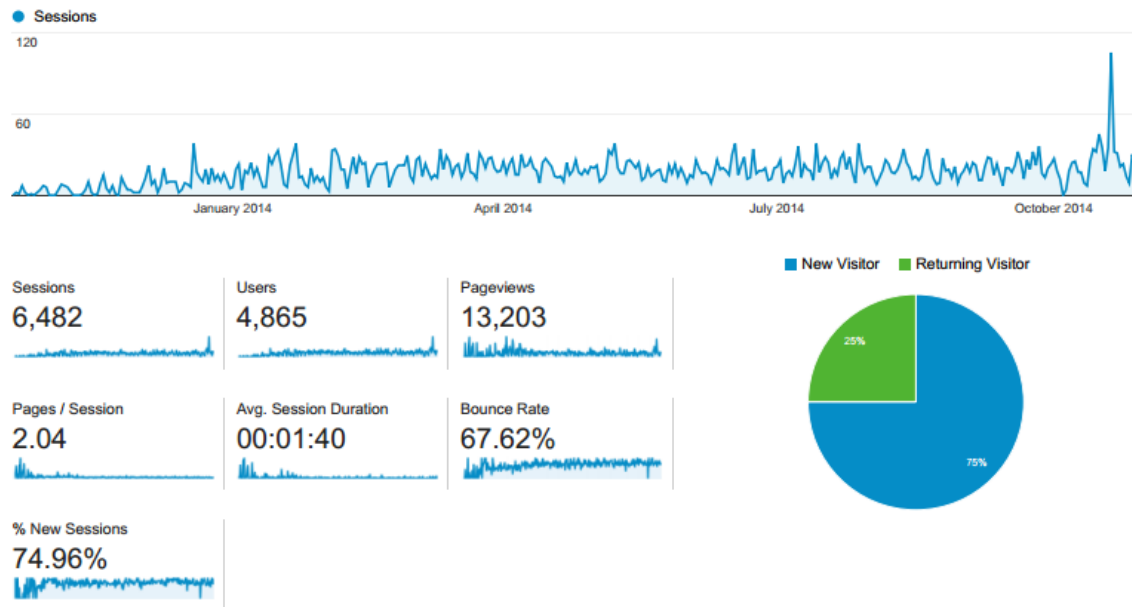
http://www.w3.org/International/its/wiki/Open_Data_Management_for_Public_Automated_Translation_Services

Pages (All pages in the wiki, including talk pages, redirects, etc.)	73
Uploaded files	27
Edit statistics	
Page edits since ITS was set up	1,104
Average edits per page	15.12
User statistics	
Registered users	35
View statistics	
Views total (Views to non-existing pages and special pages are not included)	591,908
Views per edit	536.15
Most viewed pages	
Main Page	163,124
Meetings	95,397
RulesRepository	32,102
IssuesAndProposedFeatures	30,112
ITS Implementations	24,673
ITS Translate Decorator	23,858
XLIFF Mapping	19,117
ITS Processor Interface	18,831
ITS-RDF mapping	18,570
Text Analysis serializations	18,298

6 Engagement Report: LIDER Project Portal

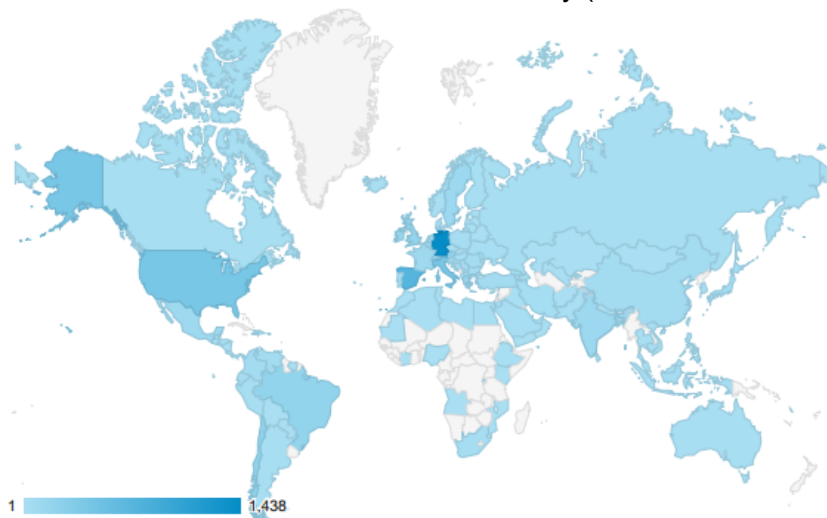
The LIDER web portal is focused on information push and as the reference site for accessing project deliverables and providing a news feed on project events. It also syndicates the Multilingual Web news feed.











General analytics (using google analytics) for the portal for the first year of operation (20 Oct 2013 to 30 Oct 2014) are given below:



Language	Sessions	% Sessions
1. en-us	1,888	29.13%
2. es	833	12.85%
3. de	770	11.88%
4. de-de	655	10.10%
5. en-gb	285	4.40%
6. es-es	281	4.34%
7. en	251	3.87%
8. pt-br	204	3.15%
9. it	151	2.33%
10. it-it	90	1.39%

A break down of these access internationally (from 121 countries) is as follows:



Country / Territory ?	Acquisition			Behavior		
	Sessions ? ↓	% New Sessions ?	New Users ?	Bounce Rate ?	Pages / Session ?	Avg. Session Duration ?
	6,482 % of Total: 100.00% (6,482)	75.01% Site Avg: 74.96% (0.06%)	4,862 % of Total: 100.06% (4,859)	67.62% Site Avg: 67.62% (0.00%)	2.04 Site Avg: 2.04 (0.00%)	00:01:40 Site Avg: 00:01:40 (0.00%)
1.  Germany	1,438 (22.18%)	65.30%	939 (19.31%)	70.45%	1.78	00:01:14
2.  Spain	742 (11.45%)	40.03%	297 (6.11%)	41.24%	4.27	00:05:31
3.  United States	427 (6.59%)	91.10%	389 (8.00%)	72.60%	1.54	00:00:56
4.  Italy	425 (6.56%)	62.82%	267 (5.49%)	63.06%	1.89	00:02:13
5.  Austria	232 (3.58%)	73.71%	171 (3.52%)	74.14%	1.77	00:00:44
6.  Chile	219 (3.38%)	99.09%	217 (4.46%)	92.69%	1.14	00:00:10
7.  Brazil	199 (3.07%)	99.50%	198 (4.07%)	95.48%	1.10	00:00:09
8.  United Kingdom	186 (2.87%)	72.04%	134 (2.76%)	54.84%	2.60	00:02:14
9.  Ireland	159 (2.45%)	55.97%	89 (1.83%)	52.20%	2.58	00:02:36
10.  Greece	153 (2.36%)	60.13%	92 (1.89%)	64.05%	1.82	00:01:53

Finally, the number of deliverable downloads at the end of LIDER phase 1 is given below:

DELIVERABLE	PUBLISHING DATE	Number of Downloads
D1.1.1 - Business use cases for the use of Linguistic Linked Data in content analytics processes - Phase I	May 12, 2014	498
D2.1.1 - Guidelines and best practices for Linguistic Linked Data-based content analytics - Phase I - v2.0	Sep 09, 2014	0
D4.1 - Community Building And Dissemination Plan v1.2	Jun 16, 2014	290
D4.3.1 - Preliminary report on LIDER Community v2.1	Jun 16, 2014	311
D4.4.1 - Project Fact Sheet, initial	Jan 03, 2014	931
D4.5 - First Roadmapping Workshop Athens v1.1	Jun 11, 2014	3058
D4.6 - Roadmapping Workshops Madrid Dublin v1.1	Sep 09, 2014	43
D5.2 - Project Portal v5.1	Jun 16, 2014	280