

Project Acronym: LeanBigData
Project Title: Ultra-Scalable and Ultra-Efficient Integrated and Visual Big Data Analytics
Project Number: 619606
Instrument: STREP
Call Identifier: ICT-2013-11

D9.2.4 Dissemination Plan and Report

Work Package:	<i>WP9 – Exploitation, Industrial Awareness, Dissemination</i>	
Due Date:	31/01/2017	
Submission Date:	31/01/2017	
Start Date of Project:	01/02/2014	
Duration of Project:	36 Months	
Organisation Responsible for Deliverable:	ICCS/NTUA	
Version:	1.0	
Status:	Final	
Author(s):	Vrettos Moulos, Achilleas Marinakis Victor Muntés, Marc Solé Luigi Romano, Ferdinando Campanile Marta Patiño Martínez Angelos Bilas Alexander J Leckey Alexandre Carvalho Ricardo Jimenez	ICCS/NTUA CA SyncLab UPM FORTH INTEL INESC UPM
Reviewer(s):	Ricardo Jiménez	
Nature:	<input checked="" type="checkbox"/> R – Report <input type="checkbox"/> P – Prototype <input type="checkbox"/> D – Demonstrator <input type="checkbox"/> O - Other	
Dissemination level:	<input checked="" type="checkbox"/> PU - Public <input type="checkbox"/> CO - Confidential, only for members of the consortium (including the Commission) <input type="checkbox"/> RE - Restricted to a group specified by the consortium (including the Commission Services)	
Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		

Revision history

Version	Date	Modified by	Comments
0.1	20/12/2016	Vrettos Moulos, Achilleas Marinakis (ICCS/NTUA)	Initial version
0.1	23/12/2016	Luigi Romano (SyncLab)	SyncLab contribution
0.1	17/01/2017	Angelos Bilas (FORTH)	FORTH contribution
0.2	18/01/2017	Marta Patiño Martínez (UPM)	UPM contribution
0.2	20/01/2017	Victor Muntés, Marc Solé (CA)	CA contribution
0.2	24/01/2017	Alexander J Leckey (INTEL)	INTEL contribution
0.2	24/01/2017	Alexandre Carvalho (INESC)	INESC contribution
1.0	30/01/2017	Vrettos Moulos, Achilleas Marinakis (ICCS/NTUA)	Final version

Copyright © 2014 LeanBigData Consortium

The LeanBigData Consortium (<http://leanbigdata.eu/>) grants third parties the right to use and distribute all or parts of this document, provided that the FIRST project and the document are properly referenced.

THIS DOCUMENT IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS DOCUMENT, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Executive Summary

The goal of work package 9 is to promote and empower the dissemination, transfer, collaboration, exploitation, assessment, and broad up-take of the LeanBigData project results to the target audience and stakeholders.

This deliverable reports on the communication and dissemination activities from all project partners during the third year. The dissemination progress is monitored by qualifying and quantifying the activities and ensuring that these efforts are sufficient to keep the project in line with the goals defined within the DoW.

This document is structured as follows: section 2 reports the promotional material that the project used during this reporting period in order to promote its results, according to the plan defined in the [2]. In sections 3 and 4 some concrete dissemination and communication actions are listed for the period M25 to M36, whereas section 5 presents an overall summary of the communication activities that took place during the three years of the project. Finally, section 6 concludes the document.

Table of Contents

Executive Summary	4
Abbreviations and acronyms	7
1. Introduction	8
1.1. Relation with other deliverables	8
2. Dissemination channels	9
2.1. Web site	9
2.1.1 Statistics from Google Analytics.....	10
2.2. Presence in Social Media	14
2.3. Promotional material	14
2.3.1 Poster.....	15
2.3.2 Promotional Video.....	16
3. Dissemination Activities	17
3.1. Publications.....	17
3.1.1 Journal Articles.....	17
3.1.2 Conference and Workshop papers.....	18
3.1.3 Patents.....	22
3.1.4 Book chapters	23
3.2. White papers	24
3.3. Workshop Organization.....	24
3.3.1 1 st Public Workshop	24
3.3.2 2 nd Public Workshop.....	25
3.3.3 3 rd Public Workshop	25
3.3.4 NTUA Hackathon	25
3.4. Training Activities	26
4. Communication Activities	28
4.1. Invited Talks	28
4.2. Project Presentations	28
4.3. Meetings and EC Events.....	30
4.4. Booths at Exhibitions.....	30
4.5. Commercial talks/meetings	32
4.6. Investor/Accelerator Pitches/Meetings	33
4.7. Workshop/Hackathon	33
4.8. Other communication activities.....	35
5. Summary of Activities throughout Project Lifecycle	37
6. Conclusion.....	38
7. References.....	39

Index of Figures

Figure 1: Project website	9
Figure 2: News section of the website	10
Figure 3: New visitors on the project web site during year 3	11
Figure 4: Top 10 countries for visitors.....	13
Figure 5: LeanBigData Logo.....	14
Figure 6: LeanBigData Poster	15
Figure 7: Technical and practical session in NTUA Hackathon.....	26
Figure 8: LeanBigData at National Centre for Scientific Research “Demokritos”.....	29
Figure 9: LeanBigData in EDF 2016.....	31
Figure 10: LeanBigData in EuCNC conference 2016.....	34
Figure 11: LeanBigData at Center for Security Studies (KEMEA).....	35
Figure 12: Official invitation from Center for Security Studies (KEMEA)	36

Index of Tables

Table 1: Publications Summary during the third year.....	17
Table 2: List of Journal Articles.....	17
Table 3: List of Conference and Workshop papers	22
Table 4: List of Patents.....	23
Table 5: List of Book chapters	24
Table 6: Training Activities	26
Table 7: Summary of Y3 Communication Activities	28
Table 8: List of Invited Talks.....	28
Table 9: List of Project Presentations	29
Table 10: List of Meetings and EC Events.....	30
Table 11: List of Booths at Exhibitions.....	30
Table 12: List of Commercial talks/ meetings	32
Table 13: List of Investor/Accelerator Pitches/Meetings.....	33
Table 14: List of Workshop/Hackathon	33
Table 15: List of Other communication activities.....	35
Table 16: Activities Summary	37

Abbreviations and acronyms

CEP	Complex Event Processing
D	Deliverable
DoW	Description of Work
EC	European Commission
ECTS	European Credit Transfer and Accumulation System
ETL	Extract Transform Load
EU	European Union
ICT	Information and Communications Technology
IoT	Internet of Things
OLAP	OnLine Analytical Processing
OLTP	OnLine Transactional Processing
PhD	Doctor of Philosophy
SQL	Simple Query Language
TBA	To Be Announced
WP	Work Package

1. Introduction

LeanBigData is an ultra-scalable and ultra-efficient big data platform integrating in one product the three main big data technologies: a novel transactional NoSQL key-value data store, a distributed complex event processing (CEP) system, and a distributed SQL database. The platform is designed to achieve scalability in a very efficient way avoiding the inefficiencies and delays introduced by current Extract-Transform-Load-based (ETL) approaches. Currently, one of the main issues in data management at enterprises and other organizations is the fact that databases are either operational (OLTP-OnLine Transactional Processing) or analytical (OLAP-OnLine Analytical Processing). This leads to a separation of the management of the operational data performed at operational databases, and the management of analytical queries performed at analytical databases or data warehouses. This separation results in having to copy the data periodically from the operational database into the data warehouse. This copy process is termed Extract-Transform-Load (ETL). ETLs are estimated to consume 75-80% of the budget for business analytics.

LeanBigData solves this issue in data management by bringing a database, LeanXcale, with the two capabilities, operational and analytical.

Another aspect in which LeanBigData innovates lies in the efficiency of the transactional processing and the storage engine. The transactional processing has been re-architected and re-implemented to be an order of magnitude more efficient than the initial version at the beginning of the project. A new storage engine, KiVi, has been architected and implemented from scratch. It is based on a new data structure to be efficient both for range queries and updates.

Another main innovation brought by LeanBigData is in the area of data streaming. Here, the goal has been to produce an efficient scalable distributed complex event processing engine

LeanBigData platform is equipped with a visualization subsystem able to report incremental visualization of results of long analytical queries and with an advanced anomaly detection and root cause analysis module. The visualization subsystem also supports efficient manipulations of visualizations and query results through hand gestures.

Four use cases have been integrated with the developed infrastructure to demonstrate the value of the LeanBigData platform and validate it.

The project is divided into nine work packages. This deliverable belongs to work package 9 and is the result of task 9.2 ("Dissemination Activities"). The aim of this task is to disseminate the results from LeanBigData through a diverse set of channels (conference and journal publications, presentations and demonstrations, press releases, web site, etc.).

1.1. Relation with other deliverables

The dissemination plan and the corresponding channels have already been defined during the first phase of the project and are clearly reported in the corresponding deliverables.

This document reports on the dissemination and collaboration activities for the third dissemination phase (M25 – M36). The measurement and evaluation of these efforts is being performed using the criteria defined in the first similar report [2]. D9.2.4 is the final version of the Communication and Dissemination Report.

2. Dissemination channels

This section offers a general description of the promotional material used for the dissemination of the LeanBigData results and reports the progress that has been made according to the dissemination plan.

2.1. Web site

The project website (<http://leanbigdata.eu/>) acts as one of the major communication vehicles for global dissemination to all specified target groups (i.e. general public, research communities, industry / business, government-related organizations).

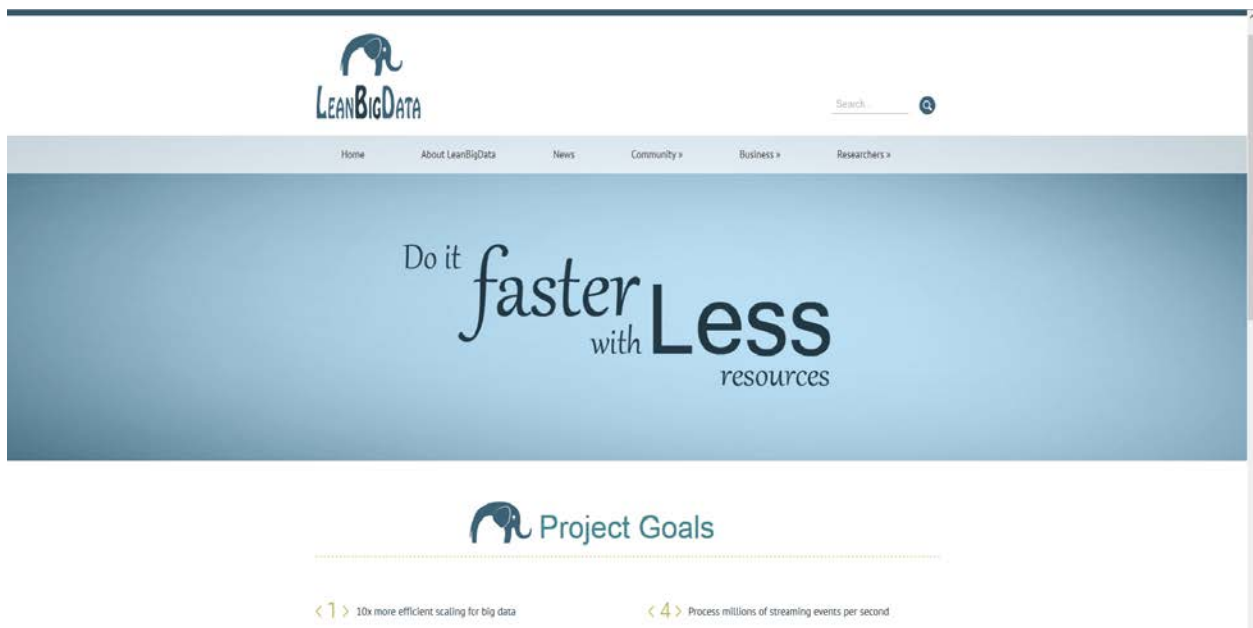
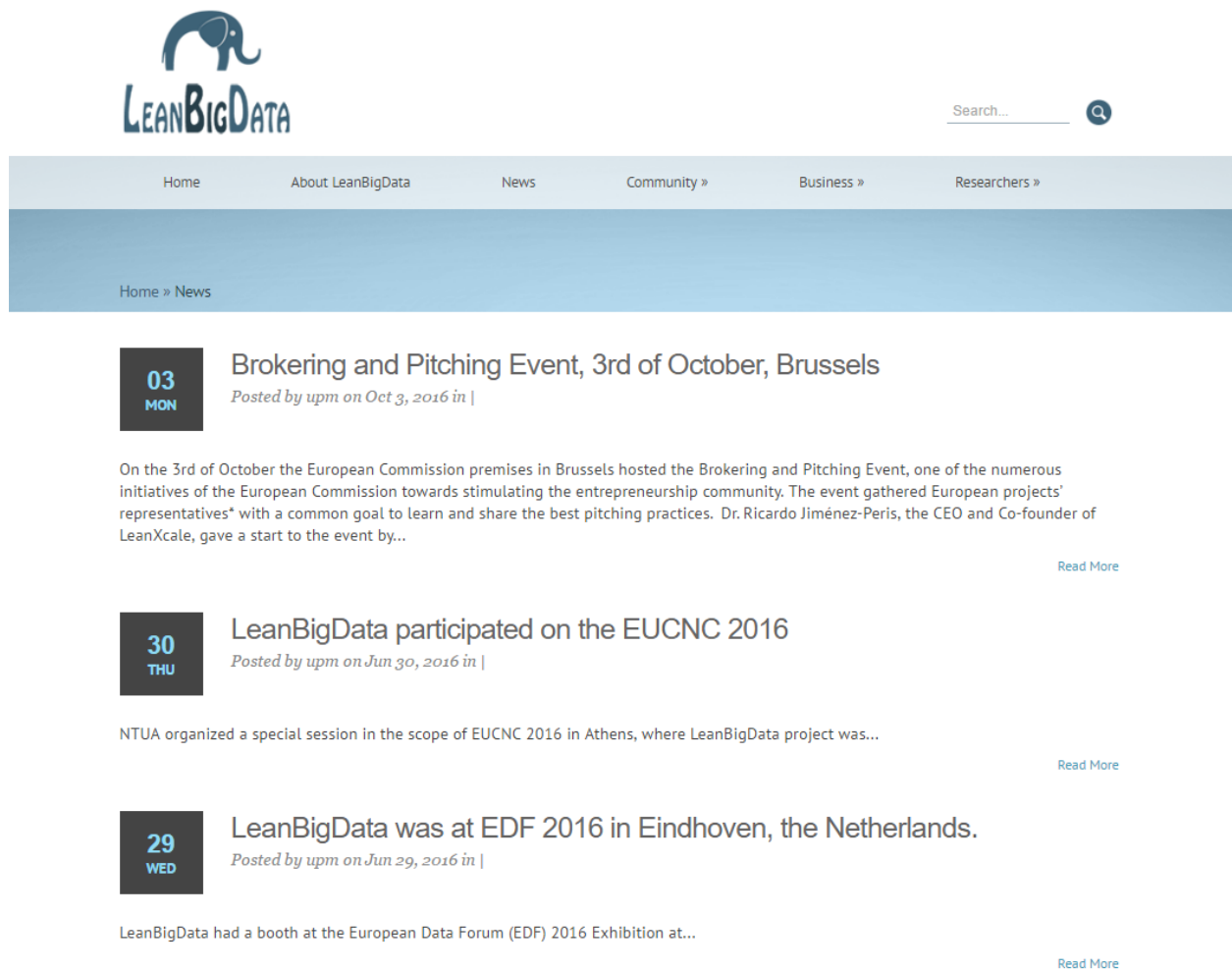


Figure 1: Project website

The content of the website includes frequently updated news regarding the presence of the project in various events and the project progress. The research section contains scientific papers and journals published by the consortium that are relevant to LeanBigData. The community corner contains the promotional material of the project, so that each visitor can access them.



The screenshot shows the LeanBigData website's news section. At the top, there is a navigation bar with links for Home, About LeanBigData, News, Community, Business, and Researchers. Below the navigation bar, there is a breadcrumb trail: Home » News. The main content area displays three news items, each with a date and a 'Read More' link.

- 03 MON** Brokering and Pitching Event, 3rd of October, Brussels
Posted by upm on Oct 3, 2016 in |
On the 3rd of October the European Commission premises in Brussels hosted the Brokering and Pitching Event, one of the numerous initiatives of the European Commission towards stimulating the entrepreneurship community. The event gathered European projects' representatives* with a common goal to learn and share the best pitching practices. Dr. Ricardo Jiménez-Peris, the CEO and Co-founder of LeanXcale, gave a start to the event by...
[Read More](#)
- 30 THU** LeanBigData participated on the EUCNC 2016
Posted by upm on Jun 30, 2016 in |
NTUA organized a special session in the scope of EUCNC 2016 in Athens, where LeanBigData project was...
[Read More](#)
- 29 WED** LeanBigData was at EDF 2016 in Eindhoven, the Netherlands.
Posted by upm on Jun 29, 2016 in |
LeanBigData had a booth at the European Data Forum (EDF) 2016 Exhibition at...
[Read More](#)

Figure 2: News section of the website

2.1.1 Statistics from Google Analytics

We use Google analytics¹ in order to collect statistics regarding the visits in our website. We analyse these statistics in order to retrieve quantitative and qualitative information so as to have better insights regarding the user experience while navigating in our website, and to see if LeanBigData really attracts visitors so that they can later return and check for updated news and information about later achievements. The results, regarding year 3, are shown in Figure 3.

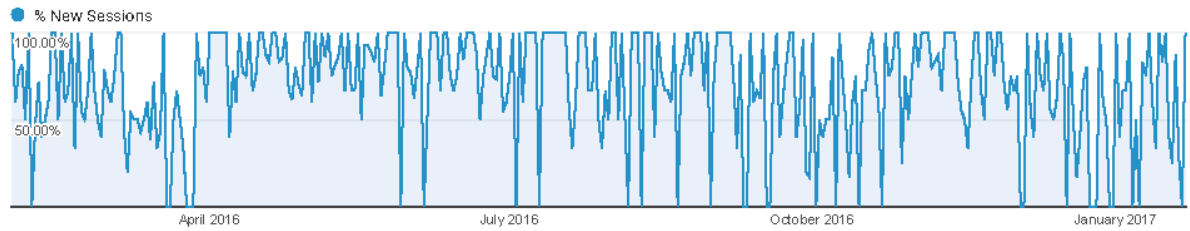
¹ <http://www.google.com/analytics/>

leanbigdata.eu

Feb 1, 2016 - Jan 22, 2017

All Users
100.00% % New Sessions

Report Tab



Date	% New Sessions	New Users
	73.02% Avg for View: 72.96% (0.08%)	1,318 % of Total: 100.08% (1,317)
1. 20160201	100.00%	7 (0.53%)
2. 20160206	100.00%	2 (0.15%)
3. 20160213	100.00%	1 (0.08%)
4. 20160214	100.00%	2 (0.15%)
5. 20160216	100.00%	7 (0.53%)
6. 20160219	100.00%	6 (0.46%)
7. 20160221	100.00%	5 (0.38%)
8. 20160225	100.00%	3 (0.23%)
9. 20160304	100.00%	2 (0.15%)
10. 20160305	100.00%	4 (0.30%)

Rows 1 - 10 of 324

© 2017 Google

Figure 3: New visitors on the project web site during year 3

The consortium enriched the project web site with updated content and used it as the main communication vehicle to disseminate the project vision and major outcomes. The goal during the final period of the project was to broaden our audience, especially towards the scientific / academic / research community (the number of publications drafted by the consortium has been increased significantly). This was noted in the Google analytics, since most of the website visitors were new sessions.

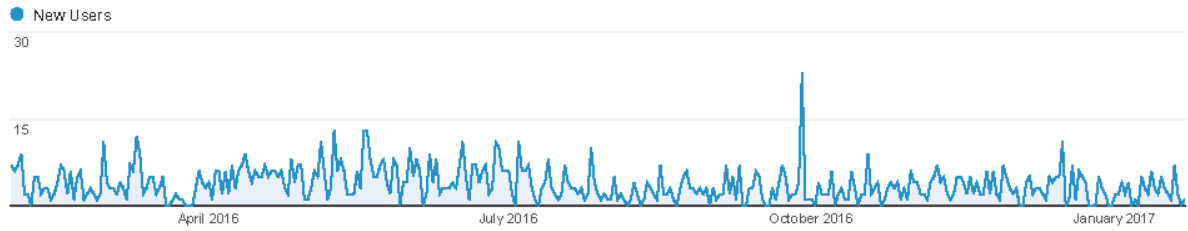
Figure 4 gives more information regarding the demography in terms of country of origin of the project web site visitors. More analytically, it is shown that the most visitors were from Spain, USA and Brazil. This means that the project triggers the interest across the globe, and it is not restricted inside the borders of the EU. This is due the fact that, apart from the regular EU events and meetings, the consortium participated in a variety of exhibition, talks and meetings across the world, in order to disseminate its vision. It is indicative that visitors coming only from USA (10,17%) and Brazil (10,09%) cover more than the 20% of the total traffic. This is explained due to the increased effort spent mainly by LeanXcale, that participated in several meetings during this year in order to attain global awareness of the project outcomes so as to attract potential customers and gain a better position in the global market.

leanbigdata.eu

Feb 1, 2016 - Jan 22, 2017

All Users
100.00% New Users

Report Tab



Country	New Users
	1,318 % of Total: 100.00% (1,317)
1. Spain	148 (11.23%)
2. United States	134 (10.17%)
3. Brazil	133 (10.09%)
4. Portugal	120 (9.10%)
5. Greece	118 (9.05%)
6. Italy	100 (7.59%)
7. India	77 (5.84%)
8. United Kingdom	71 (5.39%)
9. Germany	52 (3.95%)
10. France	46 (3.49%)

Rows 1 - 10 of 82

© 2017 Google

Figure 4: Top 10 countries for visitors

2.2. Presence in Social Media

Regarding the presence of the project in the social media, LeanBigData has an active account in Twitter and LinkedIn. However, the activity is modest, 33 followers at Twitter and 25 members at the LinkedIn group.

In Twitter, the project is present here: <https://twitter.com/leanbigdata>.

In LinkedIn, the project is present here:

<https://www.linkedin.com/grp/home?gid=6656862>.

2.3. Promotional material

Within the group of promotional material, the following tools have been considered as potent vehicles to disseminate the project messages:

- Flyers
- Posters
- Video

All the promotional material has been designed jointly with professional graphical designers in order to attract and interest visitors in public events and help to disseminate more widely the project results.

A professional logo has been designed working together with professional graphical designers to provide the project with a graphical identity:



Figure 5: LeanBigData Logo

The logo shows an elephant, one common way to represent big data nowadays, but it depicts very slim to depict that is a lean big data infrastructure, what is the core innovation to be delivered in the project.

2.3.1 Poster

A poster, unchanged as reported during the first year of the project, has been designed distilling the main message of the project in cooperation with professional graphical designers. The goal has been to create a highly attractive poster that can attract the attention of visitors in large events where the main difficulty is to attract the attention of public due to the excess of offer from many booths in exhibition floors.



Figure 6: LeanBigData Poster

The poster depicts two elephants. The one on the right depicts the current Big Data technology. As it can be seen on the elephant’s belly, it is eager on resources, as current big data technology.

Also the scale notices this showing that is on the high resource consumption side. The elephant is hairy to denote that this traditional big data technology is being rendered obsolete. The left elephant shows a slim elephant representing LeanBigData depicting that the technology being delivered by LeanBigData is “Lean”. It is also in culturist pose to denote the power of the platform. The scale reinforces the message that the platform will be low in resource usage with respect current technology.

2.3.2 Promotional Video

A promotional video has been created that serves three purposes:

- to provide a visual support for delivering a pitch about the project to broad audiences from non-technical people to experts on the topic
- to be used as background video in exhibition booths
- to be distributed through the project web site and YouTube:
<https://www.youtube.com/watch?v=R9QTmjxAdOk>

3. Dissemination Activities

This section describes all the dissemination activities that took place during the third period of the project (M25-M36).

3.1. Publications

This subsection summarizes the publication activities of the LeanBigData partners during the third reporting period. The table contains only published papers that mention the project. The following sections describe each of these activities separately.

Publications	#(M25-M36)
Journal Articles	2
Conference and Workshop papers	25
Patents	8
Book chapters	1
Total	36

Table 1: Publications Summary during the third year

3.1.1 Journal Articles

During the third period of the project the following journal articles have been published by the LeanBigData consortium.

Publication Title	Conference	Author(s) Name(s)	Publication Date	Beneficiaries
1. Transaction Management Across Data Stores.	International Journal of High Performance Computing and Networking.	Ricardo Jiménez-Peris, Marta Patiño Martínez, Iván Brondino, Valerio Vianello	July 2016 (accepted)	UPM; LeanXcale.
2. Visualization System for Monitoring Data Management Systems	Journal of Information Systems Engineering & Management, 1:4 (2016), 48 ISSN: 2468-4376	Emanuel Pinho, Alexandre Carvalho	November 7, 2016	INESC

Table 2: List of Journal Articles

3.1.2 Conference and Workshop papers

The following papers have been published as a result of LeanBigData activities:

Publication Title	Conference	Author(s) Name(s)	Publication Date	Beneficiaries
3. Snapshot Isolation for Neo4j	19th International Conference on Extending Database Technology (EDBT 2016)	Marta Patiño-Martínez, Diego Burgos-Sancho, Ricardo Jiménez-Peris, Iván Brondino, Valerio Vianello, Rohit Dhamane.	15-18 March, 2016.	UPM, LeanXcale
4. Using N-Gram Graphs for Sentiment Analysis: An Extended Study on Twitter	In IEEE Second International Conference on Big Data Computing Service and Applications	F. Aisopos, D. Tzannetos, J. Violos and T. Varvarigou	Mar 29 - Apr 1, 2016	ICCS/NTUA
5. Data Collection Framework: A Flexible and Efficient Tool for Heterogeneous Data Acquisition	DataDiversity Convergence 2016 – CLOSER 2016	L. Sgaglione, G. Papale, G. Mazzeo, G. Cerullo, P. Starace, F. Campanile	23 - 25 April, 2016	SyncLab
6. Direct Debit Frauds: A Novel Detection Approach	DataDiversity Convergence 2016 – CLOSER 2016	L. Sgaglione, G. Papale, G. Mazzeo, G. Cerullo, P. Starace, F. Campanile	23 - 25 April, 2016	SyncLab
7. PaaS-CEP: A Query Language for Complex Event Processing and Databases	DataDiversity Convergence 2016 – CLOSER 2016	Ricardo Jiménez-Peris, Valerio Vianello y Marta Patiño Martínez.	23 - 25 April, 2016	LeanXcale, UPM
8. Reducing Data Transfer in Parallel	DataDiversity Convergence 2016 –	Fábio Coelho, José Pereira, Ricardo	23 - 25 April, 2016	INESC

Processing of SQLWindow Functions	CLOSER 2016	Vilaça and Rui Oliveira.		
9. 3D Vizualization of Large Scale Data Centres	DataDiversity Convergence 2016 – CLOSER 2016	Giannis Drossis, Chryssi Birliraki, Nikolaos Patsiouras, George Margetis and Constantine Stephanidis.	23 - 25 April, 2016	FORTH
10. Design of an RDMA Communicatio n Middleware for Asynchronous Shuffling in Analytical Processing	DataDiversity Convergence 2016 – CLOSER 2016	Rui C. Gonçalves, José Pereira y Ricardo Jiménez-Peris.	23 - 25 April, 2016	INESC, LeanXcale.
11. Sentiment Analysis over politics-related big Twitter datasets	DisCoTec RTPBD 2016	F. Aisopos, V. Moulos, J. Violos, T. Varvarigou, P. Kranas, S. Stamokostas, D. Kyriazis, A. Menychtas, K. Konstanteli, G. Vafiadis, A. Evangelinou, C. Santzaridou, V. Anagnostopoulos, A. Gatzioura, N. Makrinakis and A. Papakonstantinou	June 6-9, 2016	ICCS/NTUA
12. The LeanBigData Data Collection Framework - An innovate and adaptable framework for collection and normalization of structured data	DisCoTec RTPBD 2016	Luigi Coppolino, Luigi Sgaglione, Gaetano Papale, and Ferdinando Campanile	June 6-9, 2016	SyncLab
13. Detecting Performance Degradation	DisCoTec RTPBD 2016	Dimitris Ganosis, Yannis Sfakianakis, Manolis Marazakis, and Angelos Bilas.	June 6-9, 2016	FORTH

with System Level Metrics				
14. CoherentPaas : Providing transactional support for cloud data stores.	DisCoTec RTPBD 2016	Ricardo Jiménez Peris, Marta Patiño Martínez, Ivan Brondino.	June 6-9, 2016	LeanXcale, UPM
15. STREAM-OPS: a Streaming Operator Library.	DisCoTec RTPBD 2016	Ricardo Jiménez Peris, Valerio Vianello, Marta Patiño Martínez.	June 6-9, 2016	LeanXcale, UPM
16. Targeted Advertisement case-study: a LeanBigData benchmark.	DisCoTec RTPBD 2016	Jorge Teixeira, Miguel Biscaia, Ivan Brondino, Mario Moreira.	June 6-9, 2016	Altice Labs, UPM, LeanXcale
17. Big Data Stream Clustering Algorithms Empirical Evaluation	DisCoTec RTPBD 2016	Annie Ibrahim Rana, Giovani Estrada, and Marc Solé	June 6-9, 2016	INTEL, CA
18. An RDMA Middleware for Asynchronous Multi-Stage Shuffling in Analytical Processing	The 16th IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS 2016)	Rui C. Gonçalves, José Pereira and Ricardo Jiménez-Peris.	June 6-9, 2016	LeanXcale INESC
19. A Real Time Sentiment Analysis Algorithm using Streaming Data from Twitter	Abstractions and Use Cases of converged Big Data, Telecom and IoT technologies Workshop, EuCNC 2016	V. Moulos, F. Aisopos, P. Kranas, S. Stamokostas, A. Evangelinou, M. Kardara, J. Violos, T. Varvarigou and A. Psychas	June 27-30, 2016	ICCS/NTUA
20. The CloudMdsQL Multistore System	The 2016 ACM SIGMOD/PODS Conference.	B. Kolev, C. Bondiombouy, P. Valduriez, R. Jimenez-Peris, R. Pau, J. Pereira.	Jun 26-Jul 1, 2016	LeanXcale INESC

21. A Multi-sensor Data Fusion Approach for Detecting Direct Debit Frauds	2016 International Conference on Intelligent Networking and Collaborative Systems (INCoS), Ostrawva, Czech Republic	Campanile, G. Cerullo, S. DAntonio, G. Mazzeo, G. Papale and L. Sgaglione	September 7 - 9, 2016	SyncLab
22. Polyglot Data Management Session	BOSS'16 workshop (Big Data Open Source Systems) – VLDB2016	Marta Patiño, Patrick Valduriez	September 9th, 2016	UPM
23. Anomaly Detection Guidelines for Data Streams in Big Data	3rd Intl. Conference on Soft Computing & Machine Intelligence. I SCMI 2016.	Annie Ibrahim Rana, Giovanni Estrada, Marc Solé Simó and Victor Muntés	November 23-25, 2016	INTEL, CA
24. Survey on Models and Techniques for Root-Cause Analysis	Open Access arXiv at Cornell University Library	Marc Solé, Victor Muntés, Annie Ibrahim Rana, and Giovanni Estrada	January 26th, 2017	CA, INTEL
25. Benchmarking Polystores: the CloudMdsQL Experience	IEEE Big Data – Methods to Manage Heterogeneous Big Data and Polystore Databases workshop	B. Kolev, P. Valduriez, Bondiombouy, R. Jiménez-Peris, Pau, J. Pereira.	December 5-8, 2016	INESC, LeanXcale
26. Load balancing for Key Value Data Stores.	The 20th International Conference on Extending Database Technology (EDBT 2017)	Ainhoa Azqueta Alzúaz, Ivan Brondino, Marta Patiño Martinez, Ricardo Jimenez Peris.	March 21 - 24, 2017	UPM, LeanXcale
27. Massive Data Load on	The 17th IEEE/ACM	Ainhoa Azqueta Alzúaz, Marta Patiño	May 14 - 17, 2017	UPM, LeanXcale

Distributed Database Systems over HBase	International Symposium on Cluster, Cloud and Grid Computing (CCGRID 2017)	Martinez, Ivan Brondino, Ricardo Jimenez Peris.		
---	--	---	--	--

Table 3: List of Conference and Workshop papers

3.1.3 Patents

The following patents have been submitted as a result of LeanBigData activities:

Publication Title	Patents	Author(s) Name(s)	Publication Date	Beneficiaries
28. Graph-enhanced Alerting System	Submitted to USPTO: 15/1867010	Serge Mankovski, Marc Solé Simó, Victor Muntés	Filed on 30th March 2016, not public yet.	CA Technologies
29. Method to enrich a domain knowledge-based graph describing a complex system by inferring new edges based on big data analytics to perform root cause analysis	Submitted to USPTO: 15/082,982	Victor Muntés, Marc Solé Simó, Serge Mankovski,	Filed on 28th March 2016, not public yet.	CA Technologies
30. Assisted Creation of Context Graphs	Submitted to USPTO: 15/083,046	Marc Solé Simó, Victor Muntés, Serge Mankovski,	Filed on 28th March 2016, not public yet.	CA Technologies
31. Root Cause Analysis for sequences of datacenter states	Submitted to Spanish Patent Office (not available yet)	Marc Solé Simó, David Solans, Alberto Huélamo, David Sánchez, Jaume Ferrarons, Victor Muntés	Filed on 21st December 2016, not public yet.	CA Technologies

32. Method and system for predicting future states of a Data Center	Submitted to Spanish Patent Office (not available yet)	Marc Solé Simó, David Solans, Alberto Huélamo, David Sánchez, Jaume Ferrarons, Victor Muntés	Filed on 21st December 2016, not public yet.	CA Technologies
33. Automatic data center state summarization in comprehensive way	Submitted to Spanish Patent Office (not available yet)	Marc Solé Simó, David Solans, Alberto Huélamo, David Sánchez, Jaume Ferrarons, Victor Muntés	Filed on 21st December 2016, not public yet.	CA Technologies
34. Descriptive Data Center State Comparison	Submitted to Spanish Patent Office (not available yet)	Marc Solé Simó, David Solans, Alberto Huélamo, David Sánchez, Jaume Ferrarons, Victor Muntés	Filed on 21st December 2016, not public yet.	CA Technologies
35. State information completion using graph embeddings	Submitted to Spanish Patent Office (not available yet)	Marc Solé Simó, David Solans, Alberto Huélamo, David Sánchez, Jaume Ferrarons, Victor Muntés	Filed on 21st December 2016, not public yet.	CA Technologies

Table 4: List of Patents

3.1.4 Book chapters

Publication Title	Publisher	Author(s) Name(s)	Publication Date	Beneficiaries
36. Scalable and Efficient Big Data Analytics: The LeanBigData Approach	SCITEPRESS	Ricardo Jimenez, Marta Patino, Valerio Vianello, Ivan Brondino, Ricardo Vilaca, Jorge Teixeira, Miguel Biscaia, Giannis Drossis, Damien Michel, Chryssi Birliraki, George Margetis, Antonis Argyros, Constantine	TBA	LeanXcale, Universidad Politecnica de Madrid, Altice Labs, Institute of Computer Science, Foundation for Research and Technology Hellas & Computer Science Department,

		Stephanidis, Luigi Sgaglione, Gaetano Papale, Giavanni Mazzeo, Ferdinando Campanile, Marc Sole, Victor Muntes-Mulero, David Solans, Alberto Huelamo, Pavlos Kranas, Dora Varvarigou, Vrettos Moulos, and Fotis Aisopos		University of Crete, University of Naples "Parthenope", Sync Lab srlSync Lab srl, CA Technologies, National Technical University Of Athens & ICCS
--	--	--	--	--

Table 5: List of Book chapters

3.2. White papers

Since the project has as main goal to exploit commercially most of its outcomes, most deliverables are not public to avoid the competition gaining insights on what is done before the consortium partners are prepared for doing a commercialization effort. Additionally, deliverables are quite deep technical documents with little interest for a broad audience. For this reason, a compromise has been sought that has been that every year a public white paper will be written summarizing the main outcomes of the project so far. This white paper are written with care so it is understandable to a wide audience and can attract the attention of stakeholders and potential users/clients of the platform.

For year 3 the white paper is available as deliverable 9.13 "Insight to LeanBigData".

3.3. Workshop Organization

3.3.1 1st Public Workshop

The 1st Public LeanBigData Workshop was organized on M27 of the project in Rome, Italy, on 24 April 2016. It was collocated with the conferences CLOSER 2016 [4] and IoTBD 2016 [5]. The Workshop name was "Towards Convergence of Big Data, SQL, NoSQL, NewSQL, Data streaming/CEP, OLTP and OLAP" - DataDiversityConvergence 2016 [3]. During the Workshop, 11 papers have been presented in a full day workshop.

3.3.2 2nd Public Workshop

The 2nd Public LeanBigData Workshop was organized on M29 of the project at Heraklion, Greece on 9 June 2016. It was conducted in conjunction with the event DisCoTec 2016 [6]. The Workshop's name was "Final Public Workshop from LeanBigData and CoherentPaaS" - RTPBD 2016. It was a joint Workshop with another EU project, the CoherentPaaS FP7 Project [7]. During the Workshop, 18 papers have been presented in a full day workshop.

3.3.3 3rd Public Workshop

Another workshop was organized in the top conference on databases, Very Large Data Bases (VLDB), held in New Delhi on September 9th 2016. The workshop was merged by VLDB organizers with the boss workshop and it appeared as a session within the boss workshop:

<http://vlbd2016.persistent.com/workshops.php>

<http://boss.dima.tu-berlin.de/2016/>

<http://lsd.ls.fi.upm.es/polyglot2016>

The session had 4 invited talks, including one devoted to LeanBigData delivered by Dr. Ricardo Jimenez-Peris

3.3.4 NTUA Hackathon

In collaboration with COSMOS EU project [8], ICCS/NTUA has organised a hackathon [9] in its premises in Athens on May 23rd, 2016. During the hackathon the developers had the chance to utilize the Sentiment Analysis service developed in the context of LeanBigData. Undergraduate students John Androulidakis and George Theodorakis received the second prize by creating an app called "A social application for parking". Through this app, users can search for empty parking slots in their current location. The end-users can post empty parking slots from their current position using twitter (#nodeProjectParking message) with a comment of the slot (e.g. slot for small car/ many positions for free all day). When they want to search for an empty parking slot, they can use the parking application on their phone (or computer) and get the nearest available parking slots with the comments about them. The input given is the address near which the end-user wants to park, the radius of search and the desired time window (e.g. empty slot in the last hour for more up-to-date results). Finally, the user can select the slot he parked and delete it. The user can comment about the application; the app administrators receive comments from the end-users and this is where LeanBigData kicks in, meaning that these comments are processed with the LeanBigDataSentiment in order to categorize them and improve the application in the future. The developers evaluated the Sentiment Analysis service as a useful and easy-to-use tool.



Figure 7: Technical and practical session in NTUA Hackathon

3.4. Training Activities

Table 6 gives an overview of on-going and finished teaching activities.

Type of Activities	On-Going	Finished
Lectures	8	1
PhD Dissertation	8	
Diploma Thesis	2	1
Master Thesis	2	4
Total	20	6

Table 6: Training Activities

UPM

UPM-LSD has participated in the creation of a new major in the European ICT Digital master on “Data Science” together with other top European universities such as TU Berlin, KTH Sweden, INRIA Sophi-Antipolis, etc. This is a 2-year European master (120 ECTS) with double-degree and has a minor on “Entrepreneurship”. The master aims at training entrepreneurs in hot ICT areas

where innovation activities can lead to the creation of new ICT industry. UPM is both an entry and exit node (teaches both the first and second years) and LeanXscale, the startup being created by UPM and INESC, will receive internships from the master students. In this master the course “Big Data Ecosystems” is taught by UPM-LSD.

UPM is also involved in two masters related to the project topic and UPM-LSD, the team involved in the project, participates in both. The first one is a 1-year research-oriented master “Master on Software and Systems”. Two courses are taught related to the project, “Foundations of Distributed Systems” and “New Trends in Distributed Systems”. The second master is a 2-year professional oriented master with a minor in “Entrepreneurship” and a major in “Distributed Systems”.

UPM-LSD has also 4 on-going PhD thesis and 2 master thesis have been defended on contributions of the project.

FORTH

FORTH researchers, associated with LeanBigData, contribute to the following training and teaching activities in the Department of Computer Science at the University of Crete:

- Systems programming, 2nd year undergraduate compulsory course
- Embedded systems, 4th year undergraduate elective course
- CS-527 Parallel Computer Architecture, graduate course
- 2 PhD thesis, related to storage issues
- 2 Master’s thesis, related to key value store optimizations

ICCS

ICCS contributes to the extension of the ‘Web Programming’ lecture of the under-graduate course of the school of Electrical and Computer Engineering (ECE) of the National and Technical University of Athens (NTUA) and to the extension of the ‘Fault Tolerance’ lecture of the post-graduate course of the aforementioned school. Two PhD on-going dissertations are related with the topics of LeanBigData, while two on-going diploma thesis, one finished diploma thesis and two finished master thesis are directly related with the Sentiment Analysis service that ICCS contributes to the project. Moreover, it is planned to offer a practical for students in 2017 based on the research results and prototypical implementations of LeanBigData.

CA

CA actively participates in the Aula Empresa program at Universitat Politècnica de Catalunya (UPC). In January 2016, CA gave a course on “Making IT Management Flexible, Agile and Secure”, where the company presented the Anomaly Detection and Root Cause Analysis technology they had developed in the project, and also presented the whole project its targets and its outcomes. There were 10 other courses in the program in 2016. The interest and the relevance of the event was evaluated by the students with score 4.5/5. CA’s course was ranked by students as the most interesting one.

4. Communication Activities

The following table summarizes the communication activities that took place in the third dissemination phase.

Activity	# (M25 – M36)
Invited Talks	3
Project Presentations	2
Meetings and EC Events	3
Booths at Exhibitions	10
Commercial talks/meetings	27
Investor/Accelerator Pitches/Meetings	20
Workshop/Hackathon	5
Other Communication Activities	3
Total	73

Table 7: Summary of Y3 Communication Activities

4.1. Invited Talks

No	Event Name	Event Location (City, Country)	Date	Beneficiaries
1	World Class Digital Transformation	Frankfurt, Germany	March 2016	LeanXcale
2	"Building Europe of 2020 with Novel Technological Endeavours for Smart and Complex Systems, Big Data and Future Internet" (CLOSER 2016)	Rome, Italy	April 24, 2016	UPM
3	Big Data Innovation	Frankfurt, Germany	November 2016	LeanXcale

Table 8: List of Invited Talks

4.2. Project Presentations

No	Event Name	Event Location (City, Country)	Date	Beneficiaries
1	VLDB 2016 Conference	New Delhi, India	September 2016	LeanXcale
2	Lecture at Institute of Informatics and Telecommunications, National Centre for Scientific Research "Demokritos"	Athens, Greece	18 January 2017	ICCS/NTUA

Table 9: List of Project Presentations

ICCS/NTUA presented LeanBigData (Figure 8) during a lecture that took place at the Institute of Informatics and Telecommunications, National Centre for Scientific Research "Demokritos" [14] in 18 January 2017. The audience of the lecture consisted of 16 senior and junior researchers.



Figure 8: LeanBigData at National Centre for Scientific Research "Demokritos"

4.3. Meetings and EC Events

No	Event Name	Event Location (City, Country)	Date	Beneficiaries
1	NTUA Computing research day	Athens, GR	9-Jan-2017	FORTH, keynote
2	BigStorage ITN	Mainz, GE	18-Jan-2017	FORTH, Invited talk
3	Info day Big Data	Luxembourg, Luxembourg	17-19 Jan 2017	LeanXcale

Table 10: List of Meetings and EC Events

4.4. Booths at Exhibitions

No	Event Name	Event Location (City, Country)	Date	Beneficiaries
1	Innovate Finance Global Summit	London, UK	April 2016	LeanXcale
2	Enterprise Data World	San Diego, USA	April 2016	LeanXcale
3	Collision	New Orleans, USA	April 2016	LeanXcale
4	MoneyConf	Madrid, Spain	April 2016	LeanXcale
5	European Data Forum (EDF) 2016	Eindhoven, Netherlands	29 June 2016	UPM, ICCS/NTUA, PT
6	ICSOFT conference 2016	Lisbon, Portugal	24 July 2016	ICCS/NTUA
7	DATA conference 2016	Lisbon, Portugal	24 July 2016	ICCS/NTUA
8	IoT World 2016	Barcelona, Spain	25-27 October 2016	LeanXcale
9	Web Summit 2016	Lisbon, Portugal	7-10 October 2016	LeanXcale
10	CeBit 2017	Hanover, Germany	19-23 March 2017	LeanXcale

Table 11: List of Booths at Exhibitions

The project was presented to the European Data Forum 2016 (EDF 2016) [10], which took place in Eindhoven, Netherlands 29-30 June 2016 as shown in Figure 9:



Figure 9: LeanBigData in EDF 2016

LeanBigData participated in the European Project Space at Lisbon, July 2016 during ICISOFT conference 2016 [11], which was held in conjunction with DATA conference 2016 [12].

4.5. Commercial talks/meetings

No	Event Name	Event Location (City, Country)	Date	Beneficiaries
1	Taiger meeting	Madrid, Spain	February 2016	LeanXcale
2	Qlik meeting	Madrid, Spain	February 2016	LeanXcale
3	Deutsche bank	London, UK	February 2016	LeanXcale
4	Orange meeting	Madrid, Spain	March 2016	LeanXcale
5	KPMG meeting	Madrid, Spain	March 2016	LeanXcale
6	Bull call	via Skype	March 2016	LeanXcale
7	Hopla Software meeting	Madrid, Spain	March 2016	LeanXcale
8	Santander bank meeting	Madrid, Spain	March 2016	LeanXcale
9	MapR call	via Skype	March 2016	LeanXcale
10	Odbms.org	via Skype	March 2016	LeanXcale
11	Stratio	Madrid, Spain	March 2016	LeanXcale
12	BBVA bank meeting	Madrid, Spain	May 2016	LeanXcale
13	Logitravel meeting	Madrid, Spain	June 2016	LeanXcale
14	Indizen meeting	Madrid, Spain	June 2016	LeanXcale
15	Orange meeting	Madrid, Spain	June 2016	LeanXcale
16	Mapr meeting	San Jose, USA	June 2016	LeanXcale
17	Amadeus meeting	via Skype	July 2016	LeanXcale
18	SyncLab call	via Skype	July 2016	LeanXcale
19	Analytika meeting	Madrid, Spain	July 2016	LeanXcale
20	BBVA bank meeting	Madrid, Spain	August 2016	LeanXcale
21	DAMA meeting	Madrid, Spain	September 2016	LeanXcale
22	ioFABRIC Development Center	Heraklion, Greece	Daily, throughout 2016	FORTH
23	NFQ	Madrid, Spain	October 2016	LeanXcale
24	B2IT	Barcelona, Spain	October 2016	LeanXcale
25	Neurocom	Call	November 2016	LeanXcale
26	MapR	Call	December 2016	LeanXcale
27	SERH	Madrid, Spain	January 2017	LeanXcale

Table 12: List of Commercial talks/ meetings

4.6. Investor/Accelerator Pitches/Meetings

No	Event Name	Event Location (City, Country)	Date	Beneficiaries
1	Polytech Ventures call	via Skype	February 2016	LeanXcale
2	Adara partners meeting	Madrid, Spain	February 2016	LeanXcale
3	Bullnet meeting	Madrid, Spain	February 2016	LeanXcale
4	Enterprise European Network meeting	Madrid, Spain	February 2016	LeanXcale
5	Red INNvest call	Madrid, Spain	February 2016	LeanXcale
6	Nautac call	via Skype	March 2016	LeanXcale
7	Bullnet meeting	Madrid, Spain	June 2016	LeanXcale
8	Saphire Ventures	San Jose, USA	June 2016	LeanXcale
9	INTEL Capital meeting	via Skype	August 2016	LeanXcale
10	Startup Europe Comes to Silicon Valley Investor Event	San Francisco, USA	September 2016	LeanXcale
11	EIT Digital Investor Event Berlin	Berlin, Germany	September 2016	LeanXcale
12	EIT Digital Investor Event London	London, UK	September 2016	LeanXcale
13	JP Morgan	Madrid, Spain	October 2016	LeanXcale
14	SouthSummit Pitch Finals	Madrid, Spain	October 2016	LeanXcale
15	Paua Ventures	Call	October 2016	LeanXcale
16	EC Brokering event	Brussels, Belgium	October 2016	LeanXcale
17	Schindler Ventures	Call	November 2016	LeanXcale
18	VC Meetings	Lisbon, Portugal	November 2016	LeanXcale
19	Salesforce Ventures	Lisbon, Portugal	November 2016	LeanXcale
20	EDP Carrousel Pitch Competition	Lisbon, Portugal	November 2016	LeanXcale

Table 13: List of Investor/Accelerator Pitches/Meetings

LeanXcale participated in the startup contest of the startup event SouthSummit and was selected as one of the 10 finalists in the category B2B and pitched at the SouthSummit.

4.7. Workshop/Hackathon

No	Event Name	Event Location (City, Country)	Date	Beneficiaries
1	CLOSER 2016	Rome, Italy	24 April 2016	ICCS/NTUA
2	NTUA Hackathon	Athens, Greece	23 May 2016	ICCS/NTUA
3	DisCoTec 2016	Heraklion, Greece	9 June 2016	ICCS/NTUA
4	EuCNC 2016	Athens, Greece	28 June 2016	ICCS/NTUA
5	VLDB 2016	New Delhi, India	September 2016	UPM

Table 14: List of Workshop/Hackathon

NTUA has organised a special session/workshop during the EuCNC 2016 event that was held in Athens from 28-30 June 2016. The title of the workshop was “Abstractions and Use Cases of converged Big Data, Telecom and IoT technologies” and LeanBigData was represented through the following slot:

- A Real Time Sentiment Analysis Algorithm using Streaming Data from Twitter (Fotis Aisopos, ICCS/NTUA, Greece) (Figure 10)

The workshop was organised together with CoherentPaaS FP7 project [7] and COSMOS FP7 project [8].



Figure 10: LeanBigData in EuCNC conference 2016

4.8. Other communication activities

No	Event Name	Event Location (City, Country)	Date	Beneficiaries
1	Demo at SIGMOD Conference	San Francisco, USA	June 2016	LeanXcale
2	Co-Workshop at Center for Security Studies of Hellenic Republic	Athens, Greece	19 December 2016	ICCS/NTUA
3	Meeting with CA Technologies Incubation Center to discuss about LeanBigData outcomes related to RCA	Barcelona, Spain		CA

Table 15: List of Other communication activities

ICCS/NTUA was officially invited (Figure 12) by the Center for Security Studies (KEMEA) [13], which belongs to the Greek Ministry of Interior, in order to present LeanBigData outcomes and results during a co-workshop that was organised to celebrate the 10+1 years operation of the center. Among the workshop participants were the Greek general secretary of Research and Technology and more than 500 researchers from Greece.



Figure 11: LeanBigData at Center for Security Studies (KEMEA)



Το Κέντρο Μελετών Ασφάλειας
για την επέτειο των Δέκα + 1 χρόνων λειτουργίας του,
σας προσκαλεί στην τελετή έναρξης επετειακών εκδηλώσεων
που θα πραγματοποιηθεί στις **19 Δεκεμβρίου 2016** και ώρα **10:00 π.μ.**
στις εγκαταστάσεις του Κέντρου (**Λ. Μεσογείων 96, Αθήνα**).

Στο πλαίσιο των εκδηλώσεων περιλαμβάνονται:

- Παρουσίαση του Έργου του ΚΕ.ΜΕ.Α.
- Έκθεση Ερευνητικών και Μελετητικών Αποτελεσμάτων και Τεχνολογικών Λύσεων
- Διήμερο Ενημερωτικών και Εκπαιδευτικών Δράσεων για την Εσωτερική Ασφάλεια και την Αντεγκληματική Πολιτική

Η παρουσία σας θα μας τιμήσει ιδιαίτερα.

Figure 12: Official invitation from Center for Security Studies (KEMEA)

5. Summary of Activities throughout Project Lifecycle

This section offers an overall summary of the dissemination and communication activities carried out during the whole project, so that the results can be compared between the three reporting periods. Table 16 depicts this summary:

Dissemination Channel	Year 1	Year 2	Year 3
Project Website	new	updated	updated
Flyers	new	updated	
Posters	new		
Videos	new		
Activities (events, panels etc.)	36	111	73
Publications		4	29
White Papers	new	updated	updated
Workshops			4
Hackathon			1

Table 16: Activities Summary

6. Conclusion

This deliverable reports on and evaluates the communication and dissemination activities of the LeanBigData project during the third dissemination phase (M25 – M36). The main goal from a dissemination point of view was to ensure LeanBigData visibility, which draws the attention of target groups to the project and supports the application of the project's outcomes and results. Our dissemination was based on the initial strategic planning, as reported in the previous versions of this document.

Regarding the third year of the project, as it is shown in Table 16, all partners focused on targeting the scientific community, by writing 29 publications in total, compared to 4 during year 2; the percentage change is 600%. Moreover, the consortium organised Workshops and Hackathons in order to attract more people from the academic, industrial and developers' community for better dissemination outcome. All these activities made LeanBigData known in a wider audience and it is concluded that the project's dissemination strategy was successful during the final reporting period.

7. References

- [1]. FP7 Grant Agreement - Annex II General Conditions Grant Agreement No. 611068
- [2]. LeanBigData Consortium: D9.2.2 Dissemination plan and Report, January, 2015
- [3]. DataDiversityConvergence 2016,
<http://closer.scitevents.org/DataDiversityConvergence.aspx?y=2016>
- [4]. CLOSER 2016, <http://closer.scitevents.org/?y=2016>
- [5]. IoTBD 2016, <http://www.iotbd.org/?y=2016>
- [6]. DisCoTec 2016, <http://2016.discotec.org/>
- [7]. CoherentPaaS EU Project, <http://coherentpaas.eu/>
- [8]. COSMOS EU Project, <http://iot-cosmos.eu/>
- [9]. NTUA Hackathon, <http://iot-cosmos.eu/node/1998>
- [10]. European Data Forum (EDF) 2016, <http://2016.data-forum.eu/>
- [11]. ICSOFT conference 2016, <http://www.icsoft.org/EuropeanProjectSpace.aspx?y=2016>
- [12]. DATA conference 2016,
<http://www.dataconference.org/EuropeanProjectSpace.aspx?y=2016>
- [13]. Center for Security Studies (KEMEA), Greek Ministry of Interior, <http://www.kemea.gr/en/>
- [14]. Institute of Informatics and Telecommunications, National Centre for Scientific Research "Demokritos", <https://www.iit.demokritos.gr/>