Horizon 2020 Grant Agreement No 809965
Contract start date: July 1st 2018, Duration: 30 months

LAMBDA Deliverable 2.2
Education and RTD Needs

Due date of deliverable: 31/12/2018
Actual submission date: 28/12/2018

Revision: Version 1.0

Dissemination Level		
PU	Public	x
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme, H2020-WIDESPREAD-2016-2017 Spreading Excellence and Widening Participation under grant agreement No 809965.

Author(s)	Valentina Janev, Nikola Tomašević (PUPIN), Jens Lehmann (UBO), Hajira Jabeen (UBO), Emanuel Sallinger (UOXF), Christoph Lange (Fraunhofer), Sahar Vahdati (UBO)
Contributor(s)	
Internal Reviewer(s)	Damien Graux (Fraunhofer)
Approval Date	
Remarks	

Workpackage	WP 2 Exploiting Synergies and Setting the Twinning Initiative
Responsible for WP	Fraunhofer Institute for Intelligent Analysis and Information Systems
Deliverable Lead	Institute Mihajlo Pupin (Valentina Janev)
Related Tasks	Task 2.2 SWOT Analysis, Task 2.3 Big Data Ecosystem and Major Societal

Document History and Contributions

Version	Date	Author(s)	Description
0.1	24.09.2018	Valentina Janev	First Draft
0.2	01.10.2018	Nikola Tomašević	Contribution
0.3	03.12.2018	Jens Lehmann, Hajira Jabeen	Contribution
0.4	19.12.2018	Emanuel Sallinger	Contribution
0.5	23.12.2018	Damien Graux	Review
0.6	24.12.2018	Sahar Vahdati	Contribution and Review

Executive Summary

Deliverable 2.2 is one of the key LAMBDA deliverables setting up a plan of Education, Research and Development activities for the LAMBDA consortium until the end of the project. It is based on the analysis of Big Data tools and research trends (see Deliverable 2.1), analysis of synergies and complementarities between the research and training expertise of the individual partners of the LAMBDA consortium (see Deliverable 2.3) and the needs of the main beneficiary (Institute Mihajlo Pupin) and other stakeholders from the region.

Table of Contents

Abbreviations and Acronyms

ARCH	Big Data Architecture
BDA	Distributed Big Data Analytics
HTML	Hypertext Markup Language
OBJ	LAMBDA Objective
OERs	Open Educational Resources
R&D	Research and Development
R&D&I	Research and Development and Innovation
SSE	South-East Europe Countries
SWOT	Strengths, Weakness, Opportunities and Threats
IoT	Internet of Things
SemTech	Semantic Technologies
SWOT	Strengths, Weaknesses, Opportunities and Threats.
VIS	Visualization
URI	Unified Recourse Identifier

List of Figures

List of Tables

1. Introduction

The LAMBDA project (**L**earning, **A**pplying, **M**ultiplying **B**ig **D**ata Analytics)[1] shall define a scientific strategy for stepping up and stimulating scientific excellence and innovation capacity, increasing research capacities and unlocking the research potential of the biggest and the oldest R&D Institute in the ICT area in the whole West Balkan region, turning the Institute Mihajlo Pupin[2] into a regional point of reference when it comes to multidisciplinary ICT competence related to Big Data analytics.

The preliminary Action Plan presented in the LAMBDA Grant agreement has defined ten different actions targeting four priorities

- [Open Education] Fostering efforts to create open learning resources (Big Data Analytics curriculum, educational workshops) and train the Big Data workforce (data scientists, business managers, students, and end users) in the West Balkan region;
- [Cooperation] Implementation of an Experts Exchange Program that will strengthen the partnerships and support mobility and expert exchange, knowledge and technology transfers between Serbia (PUPIN) and respectable EU research institutes and universities (via Fraunhofer, UBO and UOXF);
- [Dissemination and Outreach] Fostering cooperation with different stakeholders from the Region;
- [Exploitation and Sustainability] Exploitation of LAMBDA results in commercial projects and sustainability of the scientific activities.

Partners[3] involved in the LAMBDA project and their roles in education and training are given in Table 1.

Table 1. LAMBDA partners and main RTD roles

Short name	Partner	Organization Type	Main role
PUPIN	Institute Mihajlo Pupin, Serbia (Coordinator)	Research and Development Institute	- Maintain the LAMBDA platform, the knowledge repository for storing the learning materials, and data used for experimentation
IAIS	Fraunhofer Institute for Intelligent Analysis and Information Systems, Germany	Research and Development Institute	- Mainly involved in development of proof-of-concept solutions
UBO	Institute for Computer Science - University of Bonn, Germany	University	- Provide learning materials as Open Education Resources (OERs) via SlideWiki[4]

[1] http://www.project-lambda.org/
[2] http://www.pupin.rs/
[3] https://project-lambda.org/Partners
[4] https://slidewiki.org/

UOXF	Department of Computer Science - University of Oxford, UK	University	- Provide learning materials as Open Education Resources (OERs) via SlideWiki.org[5]

1.1 Scope

In order to clarify the strategic improvement opportunities for the Mihajlo Pupin Institute as a main beneficiary, in the WP2 framework (Task 2.2 SWOT analysis), an action was taken at the beginning of the project to analyse synergies and complementarities between the research and training expertise of the individual partners of the LAMBDA consortium. The results have been summarized in two deliverables as follows:

- Deliverable 2.2 which is a public deliverable and announces the educational and training activities of the consortium in the next two years (in a form of Education and Research, Training and Development Plan);

- Deliverable 2.3 which presents an updated picture of PUPIN's strengths, weaknesses, opportunities and threats and can be used to draw attention of the top management of partner organisations and raise awareness about the potential cooperation in future (Restricted access).

Based on the micro-level analysis and inventory of PUPIN assets and processes and macro framework (needs of Europe and the region), the goal of the SWOT analysis was also to distinguish factors for capacity building and scientific excellence, and define the key performance indicators (see also Deliverable 1.4).

1.2 Structure of the deliverable

The Deliverable 2.2 reports briefly

- on the analysis of the **Education / Research / Development Capacities and Needs** of PUPIN as main beneficiary (Section 2 and Section 3);
- on the analysis of the educational resources and capacities of LAMBDA partners for delivering lectures in Big Data domain as well as consortium plans to create resources (Big Data & Analytics curriculum and lectures, webinars) and share the resources with the community (e.g., lectures and, infrastructure), see Section 4 and also Deliverable 3.1. The *Trainers' Network* Infrastructure;
- about the organization of the Big Data Analytics Summer School in Belgrade in June 2019 (Section 5);
- about the consortium plans to implement the staff exchange programme (Section 6);
- on potential for adoption of LAMBDA outputs by other stakeholders (see Section 7 and also Deliverable 5.1).

[5] Ibid.

2. PUPIN Profile

2.1 Institute Mihajlo Pupin (PUPIN)

The Institute “Mihajlo Pupin” (PUPIN, http://www.pupin.rs/en/about-imp/profile/) is a leading Serbian R&D institution in information and communication technologies, the largest and oldest in the whole South Eastern Europe. The institute was founded in 1946 and has 465 employees, 280 of them being researchers. PUPIN’s vision is to be the leading provider of information technology and engineering solutions and services in Serbia and an important and respectable partner to EU and world-wide companies as an European-based Offshore Technology Center.

The Institute ”Mihajlo Pupin” is a member of the University of Belgrade. Research and development priorities of the Institute are closely related to the priorities defined by the Ministry of Education, Science and Technological Development of Republic of Serbia, as well as the strategies of the European Union. The overall vision of PUPIN is to address the concerns and challenges of the 21st century; to build capacities to develop large-scale complex system solutions; to remain a nucleus for nurturing Serbian innovation capacities, where individual solutions for technical and organizational problems are generated in close collaboration with industrial partners.
Taking into consideration that enabling a community-wide use of knowledge and information, and ensuring their effective utilization and universal access are prerequisites for economic, social and cultural development nowadays, PUPIN is determined to take a strategic role in a guidance of Serbian community towards a knowledge-based economy as an incubator of ICT business ideas and as an expert driving force behind development of a vivid ICT sector.
PUPIN’s vision is to be the leading provider of information technology and engineering solutions and services in Serbia and important and respectable partner to EU and world-wide companies as a European-based Offshore Technology Center.

2.2 PUPIN’s Staff

The Institute’s skilled and creative R&D workforce assembles 280 researchers mainly graduated from University of Belgrade, School of Electrical Engineering, and got their Master of Science and PhD qualifications either in Belgrade or in prominent European and global university centers. In spite of the quite large number of researchers in PUPIN, less than 30% have a scientific career and less than 10% have a career in education (5% involved in courses at universities, 5% involved in professional training activities)[6]. Less than 15% of PUPIN researchers hold a PhD degree. Majority of PUPIN researchers are involved in commercial projects and have professional career in one of the PUPIN’s key domains (energy management, traffic management, information systems, telecommunications). Although PUPIN has more than 200 professionals involved in industrial projects, the number of certified professionals (e.g. ORACLE, Microsoft, CISCO) or licenced professionals by Serbian Chamber of Engineers is quite low (currently 17 employees holds different types of licences).
The general idea for is to include all our researchers, to the bigger or lesser extent, in national and international projects.

[6] For detailed analysis, please see Deliverable 2.3.

Table 2. Examples of courses prepared by PUPIN staff and offered at universities in Serbia

University/Faculty	Course	Big Data Architect.	Semant. Technol.	Big Data Analytics	Visualization
ETF	Real-Time Systems	x		x	x
	Unmanned Autonomous Robotic Systems				
	Power Converters Power Converter Control Circuits				
SF	Information management in logistics			x	x
RAF	Artificial intelligence		x		x
	Robotics Systems				
FIT	Semantic Web Technologies		x		x
	Intelligent Software Systems		x		x

ETF - School of Electrical Engineering, University of Belgrade
SF - Faculty of Transport and Traffic Engineering
RAF - School of Computing
FIT - Faculty of Information Technology, Metropolitan University

2.3 PUPIN's Products and Services

PUPIN has an exclusive right to trademarks that are implemented in commercial projects in Serbia or abroad. The table below presents examples of projects recently completed by PUPIN:
(1) Toll collection system on the Belgrade-Niš highway, completed in 2017 (see Electronic Toll Collection, e-GO).
(2) National Distribution Dispatch center (NDDC) of the Serbian Public Electric Utility Company for managing the Operational distribution system of electric energy, commissioned in April 2016 (see VIEW4 SCADA system).
(3) The Integrated Information System developed for the Ministry of Finance, Tax Administration Belgrade, 2017

Table 3. Examples of PUPIN commercial products relevant for LAMBDA

Reference	Commercial activities relevant for LAMBDA
Toll Collection Station near Vrčin on the Belgrade-Niš highway , 2017	PUPIN has more than twenty five years of experience in design / development / implementation / maintenance of Toll Collection systems for highways, bridges, tunnels, etc. In this domain, PUPIN has references from all over the world (e.g. DR of Congo and Bosnia and Herzegovina).

	 Example: the toll collection station serves 5,600 vehicles per hour.
National Distribution Dispatch center (NDDC) of the “EPS Distribution”	 PUPIN has designed and implemented several generations of SCADA systems for large and geographically distributed managing facilities in electric power industry, water supply and irrigation, and others. Electric Power Industry of Serbia recognized Institute “Mihailo Pupin” as a highly trusted and reliable business partner, and since 2004 the Institute is actively involved in the reconstruction and maintenance, and distributed control and management of Serbian hydroelectric (VIEW H-Power) and thermoelectric (VIEW T-Power) power plants.
Ministry of Finance Tax, Administration Belgrade, 2017	PUPIN has been specializing in development, implementation, and support of advanced manufacturing and business information systems for almost 30 years. PUPIN collaborated with Ministry of Finance in the last 20 years.

Description: The Integrated Information System is based on ORACLE products.

More about PUPIN's involvement with industry can be found on PUPIN's Web site, see section **Products & Services**[7] and **References**[8].

[7] http://www.pupin.rs/en/products-services/
[8] http://www.pupin.rs/en/references/

3. Mapping of PUPIN Research, Training and Development Capacities (RTD) and Innovation Needs

3.1 IT Business Environment

Investigating drivers for innovation, the EBRD study[9] has shown that both internal and external drivers of innovation are important. In order to identify firms' capacity and ability to innovate, one has to look at both firm-level and country-level evidence.

Within the project "South East Europe IT Industry Barometer" (SEE ITIB)[10][1], analyses were conducted with the goal to assess the export capacity of ICT companies in South East Europe. The project collected basic information on ICT companies in the Western Balkan region and delivered three insight reports about the value of the national ICT market as well as their competitiveness. The project included five countries: Montenegro, Albania, Kosovo, Macedonia and Serbia. Results for Serbia have shown that Serbia has a large number of highly qualified and educated personnel in the IT area. Dominant issue regarding Serbian IT market is the fact that the industry is tied to inadequate staff training, which does not follow the needs of the IT sector in general. Companies are requiring specific educational profiles and trainings in the field of IT. Software industry and IT services industry are developed in four main segments:

1. Startup;
2. Outsourcing;
3. Development and export of original software products;
4. Development centers of large multinational companies.

SEE ITIB has also shown that most of the ICT companies from the region are software oriented offering software and IT services, while 80% of them do business with international clients within their countries. In Serbia 65% of ICT companies are members of two associations (ICT Network).

Additionally, the study entitled "IT Industry in Serbia 2015-2017" says that, since 2006, the Serbian IT industry has developed considerably. According to the 2016 data, there were close to 2,000 IT companies in our country (700 more than in 2006)[11], the number of IT employees has doubled from 10,000 in 2006 to 20,000, while the business revenue from IT services has also doubled to over 1.5 billion euros. In August 2018, 45,000 people worked in the Serbian IT sector creating software for various business sectors. Currently, most of them work for development centers of leading global technology companies like Microsoft, IBM and Intel which have already established their offices in Serbia and licenced local companies to work for them. Serbia is currently lacking 15,000 IT professionals, but with the help of the relevant authorities, the results of this sector would be even better which, in turn, would positively affect export and create added value in the IT segment.

[9] https://www.ebrd.com/downloads/research/transition/tr14c.pdf
[10] https://www.ict-net.com/language/en/ict-projects/promotion-of-regional-it/?portfolioCats=66
[11] https://www.serbianmonitor.com/en/serbia-lacks-15000-it-professionals/

3.2 PUPIN's R&D Strategy 2016-2020

The PUPIN Research and development Strategy 2016-2020 includes activities related to Big Data processing (an extract is presented in the Table 4 below).

Table 4. Overview of research topics in PUPIN Research and development Strategy 2016-2020

Topic	Subtopics
complex control systems in the **Energy domain**	- models, algorithms and software support of modern information-management systems for DSS in operational planning and management processes - applying technologies for **increased energy efficiency** - implementation of mobile (mobile / portable), energy efficient, ecological, remote controlled systems for supply of energy of small and medium power (3-100 KW) based on the use of renewable sources.
complex control systems in the **Traffic domain**, Smart cities	- Road traffic management - **Road toll collection systems** on highways, main roads, parking access control - Railway traffic management: safety, signaling and control systems
Safe and Secure Societies	- Application of ICT systems (**Cyber -Physical Systems**, Wireless Sensor Networks, Cloud Computing) - Robotic and sensory networks for remote control and monitoring using natural resources - land, water and energy in integrated, controlled agricultural production
Intelligent Information Systems, Cloud computing, and public services	- **Open and Big Data technologies (Big Data Analytics)** - Research in Cloud Computing - Intelligent Public Sector Information Management Systems and Policy Making
Telecommunication systems, **IoT and Signal processing**	- Managing QoE and Protection in the Future Internet - **Internet of Things** (IoT) - Application of digital signal processing (Systems for Predictive Maintenance of Industrial Plants - SPOIP) - "Dataflow" programming - **programming of data streams** on accelerators - reconfigurable hardware
Robotics, **Industrial production**	- **smart factories of the future and digitization of industrial production** - remote command and / or autonomous robot systems for the protection and preservation of human life and working environment (robots for firefighting, river underwater robots for the Danube and Sava rivers)

3.3 PUPIN's R&D Priorities

The research and development priorities of the Institute are closely related to the projects of the Ministry of Education, Science and Technological Development of Republic of Serbia, as well as the strategies of the European Union (Sixth and Seventh Framework Program, Interreg / CADSES, COST, Eureka, H2020, etc.). Currently, 21 projects[12] are running financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia for the period **2011-**

[12] http://www.pupin.rs/en/research-and-development-projects/national-projects/

2018. Regarding the number of the projects funded by the European Union, the Institute is the most successful Serbian research organization. Since Serbia was granted the right to participate in EU projects, the Institute has undertaken 82 projects (11 H2020, 22 FP7, 7 CIP/EIP, 3 Interreg/DANUBE, 2 Interreg - CADSES, 4 SEE, 1 IPA ADRIATIC, 4 IPA Serbia, 2 COST Actions, 1 RSEDP2, 3 TEMPUS, 1 ERASMUS+, 3 FP6, 17 bilateral).

At PUPIN projects of critical national importance have been conducted, combining systems engineering and information technology to develop process management systems (hydroelectric and thermal power plant systems), innovative solutions in the area of automation and traffic control, telecommunications and computer networks, knowledge and content technologies and applications, e-government, e-business, e-learning, etc.

Table 5. Key societal challenges in recent EU projects

Projects		Objectives
Energy Efficiency Solutions, Energy Management Systems		
Running	**RESPOND 2017-2020**	RESPOND: Integrated demand REsponse Solution towards energy POsitiveNeighbourhooDs ...Underpinned by the smart energy monitoring infrastructure, RESPOND will be able to perform **reliable energy data analytics and forecasting** in order to detect potential energy conservation opportunities, and to adapt, in real time, to the operational environment considering indoor and outdoor conditions, while retaining the requested comfort levels….
	inBETWEEN 2017-2020	inBETWEEN: ICT enabled BEhavioral change ToWards Energy EfficieNt lifestyles...inBETWEEN cloud based platform will offer advance energy services that allows users to integrate their building's connected devices and systems with **advance energy analytics and optimisation services** to create a comprehensive recommendation and feedback solution which will facilitate further the behaviour change towards more energy and cost efficient daily routines.
	IDEAS, 2018-2021	**IDEAS –** Novel building **I**ntegration **D**esign for increased **E**fficiencies in **A**dvanced climatically tunable renewable energy **S**ystems ...The research aim is to create a novel low-cost building integrated renewable energy system maximizing the output tuneable for different climatic conditions through novel luminescent and geometric concentrator techniques leading to current solar system efficiencies being exceeded electrically. ...This building integrated RES will use **advanced control techniques** to maximize performance and electrical and thermal/cooling self-sufficiency in the building.
	REACT, 2018-2021	**REACT – Renewable Energy for self-sustAinable island CommuniTies** REACT will demonstrate the potential of large-scale deployment of RES and storage assets on geographical islands to bring economic benefits, contribute to the decarbonisation of local energy systems, reduce greenhouse gas emissions (GHG) and improve air quality…..REACT will leverage on energy production and consumption modelling, **grid operation fault detection and diagnostics**, **multi-carrier supply optimization** and optimal energy dispatching, while enabling synergies between different energy networks and micro-grids of island.
Safe and Secure Societies		
Completed	**SPARTACUS, 2013-2016**	**SPARTACUS –** Satellite Asset Tracking for Supporting Emergency Management in Crisis Operations ...develop **Galileo-ready satellite-based applications**...solutions for Location Awareness in emergency management that can be deployed during disaster, replacing the traditional use of terrestrial networks… See for example the FLARE system [13] for real-time tracking and situational awareness of first responders in emergency situations, developed by PUPIN.

[13] http://www.sofia.rs/docs/tr/2015/SOFIA-TR8.pdf

	EMILI, 2010-2013	**EMILI** - Emergency Management in Large Infrastructures ...research in **operational and data management, complex event processing**, reactivity, knowledge representation, … See for example the Airport Use Case Demo of the Recommendation and Decision Support System developed by PUPIN[14].
Intelligent Information Systems and Public services		
Completed	LOD2, 2011-2014	**LOD2** – Creating Knowledge out of interlinked Data...LOD2 aims at developing: 1. tools and methodologies for exposing and managing **very large amounts of structured information**, 2. a testbed and bootstrap network of high-quality multi-domain, **multi-lingual ontologies**. 3. **machine learning algorithms** for (semi-)automatically enriching, interlinking and fusing data from the Web. 4. adaptive tools for exploring, searching, browsing, and authoring of Linked Data. The project showcases the benefits for media & publishing, corporate data intranets and eGovernment. See for example the results of the PUPIN team • ROZETA - Multilingual and multifunctional Natural Language Processing (NLP) and Linked Data tool, 2014 (Demo, Video,Leaflet) • LOD Statistical Workbench, 2014 (Demo, videti LinkedData Stack, Leaflet, PDF)
	GeoKnow	**GeoKnow –** Making the Web an Exploratory for Geospatial Knowledge...addresses challenges in the area of intelligent information management: the exploitation of the Web as a platform for **geospatial knowledge integration** as well as for exploration of geographic information…..creation and maintenance of qualitative geospatial information ...Tools and methodologies for mapping and exposing existing structured geospatial information on the web of data, considering comprehensive and qualitative ontologies and efficient spatial indexing. See for example the results of the PUPIN team • Exploratory Spatio-Temporal Analysis of Linked Data, 2015 (Demo, Video, GitHub, PDF) • RDF Data Cube Validation Tool, 2015 (Demo, GitHub, html, PDF) • Geospatial-semantic Exploration on the Move, 2015 (GitHub, Video 1, Video 2, Leaflet)

LAMBDA proof-of-concept development activities shall be aligned with the running EU projects and shall sustain the development of innovative solutions started in past EU projects.

3.4 Examples of Innovative Scenarios based on PUPIN's products

3.4.1 Energy Domain Scenarios

ENERGY USE CASE
Energy supply chain, starting from electricity production, over the transmission and distribution, to the consumption, represents a data intensive environment, generating a high volume of data (coming from circuit elements, smart metering, etc.) and making it difficult for supervision and control. Therefore, tasks of ensuring a reliable operation of power grid and delivering efficient user driven services pose a number of challenges, such as: **Challenges:** • Improved stability and flexibility in the distribution grids with large share of renewables, providing more

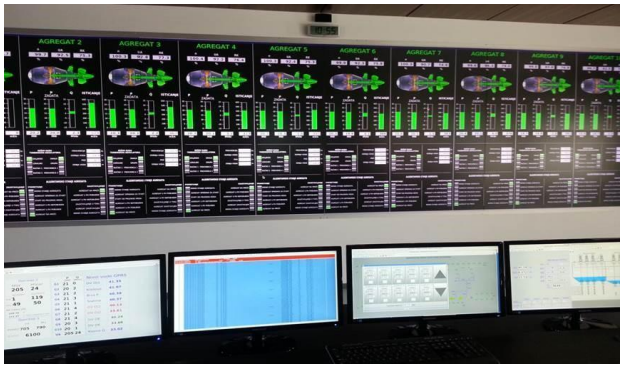
[14] Mijović, V., Tomašević, N., Janev, V. et al. J. Syst. Sci. Syst. Eng. (2018). https://doi.org/10.1007/s11518-018-5393-5

efficient operation and becoming adaptable, autonomous and self-healing,
- Advanced energy management & control systems, and reliable energy production & consumption forecast to provide efficient smart grid, user oriented services,
- Advanced tools for power grid operation analytics, such as fault detection and diagnostics, predictive maintenance, to ensure efficient and optimized power distribution, reliable grid operation and network load balancing (in particular of MV and LV grids),
- Power grid infrastructure planning tools to support next generation technologies enabling smart grids with increased share of renewables, and to provide highly modular and adaptable power grids.

PUPIN performs supervision and control of entire Serbian energy supply chain, covering electricity production (thermal, hydro and PV power plants), transmission and distribution (and consequently electricity consumption). Furthermore, PUPIN was responsible for deployment and implementation of national dispatching centers where all relevant data acquired from Serbian power grid are monitored.

Serbian power transmission and distribution network is composed of power substations (where electricity transformation from high to medium, or medium to low voltage (and vice versa) is carried out) connected by power lines for conducting the electrical energy. Power substation are equipped with smart metering equipment and intelligent electrical devices (IEDs) which, apart from raw data (directly acquired from meters and other circuit elements), send out preprocessed data as well.

All signals are acquired periodically (in regular time intervals) or upon specified status change (if certain circuit element changes its position to on/off). Data format of acquired signals is composed of signal ID, measured value, time stamp, quality flag (which indicates alarm, value is out of range, invalid data source, communication failure, etc.) and other related parameters.

National energy dispatching center (Implemented by PUPIN)	**Hydroelectric power plant "Djerdap" Control Room (Implemented by PUPIN)**
	

For supervision and control of Serbian power grid, PUPIN developed and deployed SCADA View4 system which, apart from data acquisition, also performs real-time processing of raw data (such as verification that measured values are within predefined range, signal aggregation and filtering, alarm prioritization, hierarchically aggregation of alarms, etc.). All data, both raw data acquired in real-time (e.g. directly sent out from meters) and their preprocessed derivatives are logged in designated SCADA archives.

Available datasets for experimentation:

In total, SCADA monitors around 100,000 analog and digital signals, which most of them are logged every 10 milliseconds. Currently, SCADA archives contain power grid data logged in interval of last 10 years, which amounts to petabyte order of magnitude. These archived data bring in a huge potential for performing precious intelligent analytics that could be undertaken to forecast the power production and consumption, system troubleshooting, etc. as specified in the following use case scenarios. Deployment of BigDataEurope (BDE) Software Stack will ensure interoperability and integration of high-volume heterogeneous data sources (such as power substations, smart metering, and external data sources) providing efficient smart grid and energy management services to the end consumers, as well as advanced power grid operation and planning tools to the power grid operators.

- Dataset of power data records coming from the MV/LV power substations are regularly archived in SCADA View4 databases
- Approximately 100,000 analog and digital signals (on power consumption, energy conversion, voltage transformation, etc.) are recorded and logged every 10 milliseconds
- SCADA archives contain power grid data logged in interval of last 10 years (petabyte order of magnitude)

Specific scenarios	
Provision of smart grid services – Advanced energy management,	**Enabling intelligent smart grid services:** To provide advanced smart grid,

energy production & consumption forecast	user oriented services, **integration with high volume heterogeneous smart metering data** (coming both from grid side, e.g. placed in power substations and from user side, e.g. installed in homes, buildings) is a prerequisite. To specify, suggest and deliver adequate services to the end users (i.e. energy consumers) with respect to their requirements and power grid status, various energy data analytics should be applied by DSO and gird operators. One of them is to provide precise **short- and long-term energy production and consumption forecast**. In order to deliver such energy production and consumption forecast analytics, **historical energy production data** (such as from thermal and hydro power plants, RET power plants) and **historical consumption data** (based on smart metering at consumer premises or LV/MV power substations) have to be taken into account. Such precise production and demand forecasts will **further enable advanced energy management and control services** in terms of energy conservation measures, demand side management strategies, demand response activities, etc. As main challenge in provisioning of advanced smart grid services is related to the **integration and interoperability of high volume heterogeneous data sources** (from smart metering and substation metering equipment) as well to the adequate processing of the acquired data. Furthermore, making this data **interoperable (based on Linked Data API) and interlinked with other data sources** (such as weather data for RET production analysis, number of inhabitants per home units, etc.) is essential for providing efficient user tailored services (energy conservation action suggestion, comparison with other consumers of the same type, etc.).
Power grid operation analysis - Fault detection and diagnostics, predictive maintenance	**Ensuring reliable operation of power grid:** In order to provide efficient and optimized power distribution, reliable grid operation, network load balancing (in particular of MV and LV grids), adequate **tools for real-time power grid operation analysis** are required (including power flow, energy conversion, voltage transformation, etc.). Apart from the analysis of power grid operating in regular conditions, it is important to provide adequate support to the **analysis of power grids in exceptional and alarm situations** (e.g. in case of power outage, faulty elements) including system troubleshooting, fault detection and diagnostics. To provide such advanced analytics, **real-time integration and big data analysis performed upon the high volume data streams** coming from metering devices and power grid elements (e.g. switches, transformers, etc.) is necessary. In this way, intelligent alarm and raw data processing will be deployed in order to discover the cause of failure (e.g. why circuit breaker went off, the reason of communication failure, etc.) and potential consequences (such as fault avalanches). Moreover, by integrating with such real-time heterogeneous data from the power grid, **provision of autonomous and self-healing grids** will be enabled. In this regard, predictive maintenance algorithms (by power circuit element operation analysis) will provide **reliable prediction of maintenance or replacement activities** with respect to the operational life cycle of electrical switching or other circuit elements. The main challenge in this specific scenario will be to provide **interoperability and integration with heterogeneous and high volume data sources** (at power substation level in MV and LV grids) using Linked Data principles. In addition, integration with power grid operation analytics should be ensured **leveraging upon Linked Data API**. As a result, suitable and intuitive way for visualization of the acquired, processed and derived data will be provided to the DSO operators thus improving the overall **real-time system awareness**.
Power grid infrastructure planning – Power system design and reconfiguration	**Providing an optimal design of power grid:** To support next generation technologies enabling smart grids with increased share of renewables, it is necessary to provide **highly modular and adaptable power grids**. In addition, to cope with increasing trend in energy demand while meeting H2020 goals, adequate tools for **off-line analysis of power system optimal design** should be deployed. These analytical tools should also incorporate

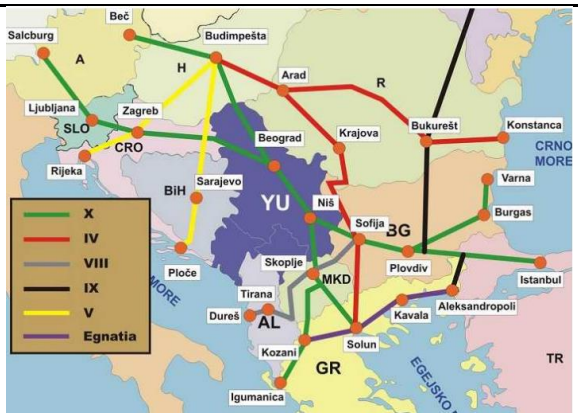
	allocation of optimal reconfiguration of power grid elements to provide reliable and flexible operation as an answer to the changing operational conditions (e.g. energy demand peaks). Tools for planning and reconfiguration of power distribution network consider planning of power stations infrastructure, number and capacity of power lines, etc. To provide such advanced grid capabilities, **integration with historical power grid data, archives of detected alarms and other relevant operation data** (such as from smart metering, consumption data, etc.) is necessary. Therefore, the main challenge in this scenario is to provide digested input to the **batch-processing, big data analytics for power grid infrastructure planning**. To achieve this, as mentioned, integration and **interoperability with heterogeneous data sources using Linked Data principles** (such as yearly data archived from power substations, smart metering devices, but also static grid infrastructure data) is a prerequisite. This should also include the **integration of both production and consumption data** in order to ensure that production side meets the energy requirements and demand side of the power grid.

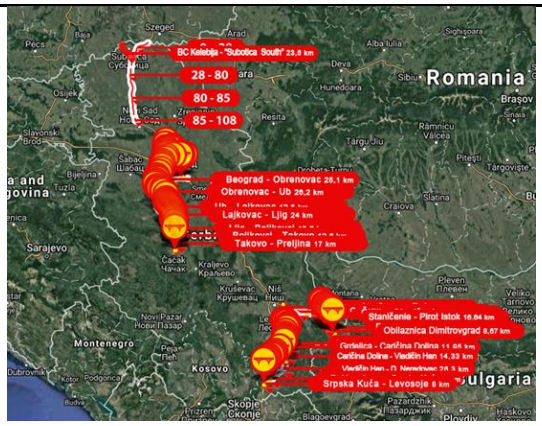
3.4.2 Traffic Domain Scenarios

TRAFFIC MANAGEMENT USE CASE
Highways and motorways control systems generate a high volume of data that is relevant for a number of stakeholders from the traffic and environmental departments, to transport providers, citizens and the police. **Challenges:** • EU transport networks should be more efficient[15,16] • Interoperability of tolling services on the entire European Union road network (European electronic road toll systems introduced at local and national levels from the early 1990s onwards were, and generally still are, non-interoperable)[17] • users of the EU highways and motorways need to be better informed, see Smart Mobility opportunities in the Netherlands[18] • An efficient and effective approach for assessment and management of air pollution due to road transport is necessary to improve ambient air quality. Since 1980s, PUPIN is involved in development and implementation of complex supervisory and control systems in the field of transport including (a) toll collection and motorway and highway traffic control systems, and (b) urban traffic control and management. Figure below points to (a) the European corridors, and (b) the **Corridor X** that is managed in Serbia by the public enterprise "Roads of Serbia"[19] using a control system provided by PUPIN.

European corridors	**Corridor X in Serbia**

[15] http://www.transport-research.info
[16] http://www.transport-research.info/sites/default/files/TRIP_Transport_infrastructures_0.pdf
[17] http://ec.europa.eu/transport/sites/transport/files/media/publications/doc/2011-eets-european-electronic-toll-service_en.pdf
[18] https://www.tno.nl/media/7613/magazine-smart-mobility.pdf
[19] http://www.putevi-srbije.rs/index.php/en/





The toll collection system, http://static.pupin.rs/2016/02/Toll-Collection-System-EN.pdf, is hierarchically structured; it is fully modular, based on PC technology and up-to-date real time operating systems, relational data base system and dedicated encryption of data transmission. Toll line controllers are based on industrial PC-technology and dedicated electronic interface boards. Toll plaza subsystem is the supervisory system for all line controllers. It collects all the data from lane controllers including financial transactions, digital images of vehicles, technical malfunctions, line operator's actions and failures. All data concerning toll collection process and equipment status are permanently collected from the plaza computers and stored in a central system database. Main Control center is connected through an optical communication link with Plaza Control centers. Also, center is constantly exchanging data with various institutions such as: banks, insurance companies, institutions that handle credit and debit cards, RF tags vendors, etc. through a computer network.

The toll collection system comprises also features concerning vehicle detection and classification, license plate recognition and microwave based dedicated short range communications (DSRC).

Data analytics is based on data warehouse architecture enabling optimal performances in near real-time for statistical and historical analysis of large data volumes. Reporting is based on optimized data structures, allowing both pre-defined (standardized) reports as well as ad-hoc (dynamic) reports efficient generation utilizing Oracle BI platform.

By coupling the BigDataEurope (BDE) Software Stack with the PUPIN toll collection system and integrating transport data with data assets from other sectors (environment, tourism, etc) innovative services can be implemented for the benefits of motorway/highway users, public authorities and society. Currently for the analysis of traffic flow in Serbia, traffic engineers calculate the indicators on an annual basis (such as annual average daily traffic, average speed, average speeding, etc.).

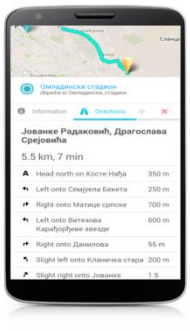
Available datasets for experimentation:

- Example dataset from the central system database before preparing the hourly / daily statistics
- Approximately 1.5 billion records of daily statistics on traffic volumes, traffic structure, vehicle speeds, etc. are generated on yearly basis.
- Summary reports for the period 2005-2015 are public in PDF format, http://www.putevi-srbije.rs/index.php/en/traffic-counting
- Growth rate of yearly average daily traffic 2015/2014 was 6.3%

Re-use of tools:

- BigDataEurope (BDE) Software Stack
- Geospatial-semantic Exploration on the Move, 2015 (GitHub, Video 1, Video 2)

Specific scenarios	
Smart mobility - Developing next generation journey planner applications	**Improved search and retrieval of traffic data:** Route planners need to be provided with a lot of data from various sources: geographical data, speed limits, road blocks, time schedules, real-time vehicle locations, etc. Making the traffic counting data interoperable e.g. as a Linked Data API (**Corridor 10 Linked Data API**) enables easy integration and using data from Serbia in other open source projects such as OpenTripPlanner, as well as developing next generation journey planner applications. **Motive-based search service** can be implemented based on

	linking the **Corridor 10** infrastructure and tourist information available in Linked Data format (e.g. from DBpedia). Challenges in this scenario are related to **interlinking of static data sources** (different input formats, granularity, representations of the road infrastructure, etc) **with dynamic data** from the toll collection system. **Resource Details in GEM: a) Directions view; b) motive-based search**  
Situational awareness - Predicting and preventing road traffic congestion	**Predicting and preventing road traffic congestion:** The goal of the scenario is to improve the congestion diagnosis and to enable traffic managers to proactively manage the traffic and organize the activities at toll collection stations before congestion is reached. **Traffic status service** based on real-time data can provide useful information (traffic reports, traffic alerts) for predicting road traffic congestion will be developed. Besides traffic counting data, other data sources can be consulted such as weather data (temperature, rainfall, fog, …), video surveillance and external data (e.g. triggers via external APIs). Challenges in this scenario are related **to querying and reasoning over heterogeneous streams**. The novelty of the scenario lies in the ability of the BDE tools to ingest highly heterogeneous real-time data and perform various types of semantic inferences i.e., analysis, diagnosis, exploration and prediction. **Traffic status prediction service** implemented using Linked Data principles (standard vocabularies, semantic stream enrichment, semantic association rules) will be interoperable with similar cross-border services.
Better policy-making based on integrated and comprehensive analysis of road traffic data for different purposes (e.g. Linking transport, environment and sustainability)	**Strategic Environmental Impact Assessment:** The goal of the scenario is to study the **environmental impact** and the effect of highways on adjacent flora, fauna, air, soil, water, humans, landscape, cultural heritage, etc based on **historical and real-time analysis**. Passive pollution monitoring involves collecting data about diffusion of air pollutants e.g. emission estimates based on traffic counting. Passive pollution monitoring has been used to determine trends in long term pollution levels. For road traffic pollution monitoring and visualization, **integration of high volume (historical) traffic data** with other parameters such as **vehicle emission factors, background pollution data, metrology data, and road topography** is needed. The service, which development will be aligned with the Technical guidelines environmental impact assessment in road sector in Serbia[20], can be used **for environment protection planning** e.g. optimization of construction, maintenance of interurban roads as foundation of total social-economic development. Hourly / daily traffic data can be extracted from the toll collection system. Historical data on average annual daily traffic referring to public roads of the I and II category in the Republic of Serbia is open and provided by the "Roads of Serbia" public enterprise, http://www.putevi-srbije.rs/index.php/en/.

[20] Sector for strategy, designing and development, Department for environmental protection, http://www.putevi-srbije.rs/images/pdf/strategija/tehnicko_uputstvo_procena_uticaja_na_zivotnu_sredinu_eng.pdf

4. Education and RTD Capacities of the LAMBDA Consortium - Big Data Analytics Lectures

Knowledge transfer by internationally-leading counterparts to PUPIN will be conducted as follows.

- UBO and UOXF will be mainly involved in preparing teaching materials (see OBJ 2.A1 Grant Agreement) and conducting the training at PUPIN premises (see OBJ 2.A3 Grant Agreement);
- IAIS and UBO will be involved in training PUPIN staff for elaborating IT solutions for concrete problems formulated by industry experts (see OBJ 2.A2 Grant Agreement).

Taking into consideration that both aspects (education) and (piloting) are important for PUPIN, the Big Data Analytics Summer School will be organized as an event that brings together researchers and professionals to discuss state-of-the-art in Big Data research and applications. More information about the organization of the Big Data Analytics School is given in next Section.

4.1 Analysis of available UBO/IAIS lectures

The University of Bonn (UBO) is working on the cutting edge technologies related to Big Data, Intelligent Analysis and Information Systems. The concerned team at the Smart Data Analytics (SDA) group is active in specializing applied research in intelligent data and knowledge analysis and teaching activities of the relevant topics. In order to share their knowledge and expertise, the following six lectures have been planned to be delivered by UBO team.

4.1.1 Introduction to Big Data Architecture

[ARCH-Lecture-1] This lecture will cover the existing advanced Big Data architectures following a bottom-up approach. In this lecture, the important knowledge to design and architect scalable solutions for challenging problems will be introduced. The primary components in the architecture of such systems and their architectures will be presented and discussed including “inter alia distributed kernels” and cluster managers, distributed file systems and storage systems.

4.1.2 Big Data Solutions in Practical Use-cases

[ARCH-Lecture-2] This lecture focuses on architecting Big Data solution. We will discuss the role and importance of the components in realizing system architectures. The participants will be introduced to unique problem characteristics that drive Big Data and the unending technology options to solve them. The application of the introduced concepts and components will be discussed in real-world example of practical use-cases.

4.1.3 Distributed Big Data Frameworks

[ARCH-Lecture-3] The “processing frameworks” are one of the most essential components of a Big Data systems. There are three categories of such frameworks namely: Batch-only frameworks (Hadoop), Stream-only frameworks (Storm, Samza), and Hybrid frameworks (Spark, Hive and Flink). In this lecture, we will introduce them and cover one of the major Big Data frameworks, Apache Spark. We will cover Spark fundamentals and the model of “Resilient Distributed Datasets (**RDDs**)” that are used in Spark to implement in-memory batch computation. Furthermore, essential parts of the important practical techniques will be introduced such as Hadoop Distributed File System for the data resiliency, and the "lineage" property of “Directed Acyclic Graphs (DAG)” to achieve resilience for the computation resiliency, or use of catalyst for code optimization.

4.1.4 Distributed Big Data Libraries

[SBDA-Lecture-1] In the practical level, the Big Data frameworks use different APIs for graph computations and graph processing. In this lecture, the important libraries built on top of Apache Spark will be covered. These include SparkSQL, GraphX and MLlib. The audience will learn to build scalable algorithms in Spark using Scala.

4.1.5 Distributed Semantic Analytics I

[SBDA-Lecture-2] This module will cover the needs and challenges of distributed analytics and then dive into the details of scalable semantic analytics stack (SANSA) used to perform scalable analytics for knowledge graphs. It will cover different SANSA layers and the underlying principles to achieve scalability for knowledge graph processing.

4.1.6 Distributed Semantic Analytics II

[SBDA-Lecture-3] This module will cover the setup, APIs and different layers of SANSA. At the end of this module, the audience will be able to execute examples and create programs that use SANSA APIs. The final part of this lecture is planned to be an interactive session to wrap up the introduced concepts and present attendees some open research questions which are nowadays studied by the community.

4.2 Analysis of UBO/IAIS tools

SANSA Stack: A major research challenge is to perform scalable analysis of large-scale knowledge graphs to facilitate applications like link prediction, knowledge base completion and reasoning. Analytics methods which exploit expressive structures usually do not scale well to very large knowledge bases, and most analytics approaches which do scale horizontally (i.e., can be executed in a distributed environment) work on simple feature-vector-based input. The Semantic Analytics Stack (SANSA) supports expressive and scalable semantic analytics by providing functionality for distributed computing on RDF data. SANSA uses Spark and Flink which offer fault-tolerant, highly available and scalable approaches to process massive sized datasets efficiently. SANSA provides the facilities for Semantic data representation, Querying, Inference, and Analytics.
SANSA-Stack's core is a processing data flow engine that provides data distribution and fault tolerance for distributed computations over RDF large-scale datasets. SANSA includes **several libraries** for creating applications:

1. **Read / Write RDF / OWL library** for RDF/OWL operations,
2. **Querying library** support a query language on top of distributed RDF/OWL library,
3. **Inference library** implements rule-based reasoning on RDF/OWL data,
4. **ML- Machine Learning core library**.

In the LAMBDA framework, **SANSA Stack will be tested in different scenarios with PUPIN data.** Staff exchange will be organized in February 2019 and May 2019.

4.3 Analysis of available UOXF lectures

The University of Oxford (UOXF) is working on cutting edge technologies related to Big Data and analytics. The concerned team at the VADA ("Value Added Data Systems") group is active in research and teaching activities with regard to these topics. In order to share their knowledge and expertise, the following lectures have been planned to be delivered by the Oxford team:

4.3.1 Introduction to Knowledge Graphs

[EKGs-Lecture-1] This module will introduce the topic of Knowledge Graphs. We will cover what a Knowledge Graph is, the similarities and differences between “world” Knowledge Graphs and Enterprise Knowledge Graphs, as well as theory and practice in the area. In particular, we will discuss the Vadalog Knowledge Graph Management System developed at the University of Oxford.

4.3.2 Reasoning in Knowledge Graphs

[EKGs-Lecture-2] This module will discuss reasoning in Knowledge Graphs. Reasoning is essential to gain value from Knowledge Graphs by deriving insights, and making available new implicit data from existing data. We will cover the theory and practice of reasoning in Knowledge Graphs, and provide a number of easily accessible examples based on Oxford’s Vadalog system.

4.3.3 Extraction for Knowledge Graphs

[EKGs-Lecture-3] This module will discuss the topic of extraction for Knowledge Graphs. We will focus on web data extraction in this module. Web data extraction is essential to make information available on the web accessible and usable by Knowledge Graphs. We provide a thorough introduction to the topic. This will feature both Oxford’s Vadalog and OXPath systems.

4.2 Analysis of UOXF tools

Vadalog System: The Vadalog system is a Datalog-based system for performing complex logic reasoning tasks, such as those required in advanced knowledge graphs. Over the past years, there has been a resurgence of Datalog-based systems in the database community as well as in industry. In this context, it has been recognized that to handle the complex knowledge-based scenarios encountered today, such as reasoning overlarge knowledge graphs, Datalog has to be extended with features such as existential quantification. Yet, Datalog-based reasoning in the presence of existential quantification is in general undecidable. Many efforts have been made to define decidable fragments. Warded Datalog+/- is a very promising one, as it captures PTIME complexity while allowing ontological reasoning. Yet so far, no implementation of Warded Datalog+/- was available.
The Vadalog system is Oxford’s contribution to the VADA research programme, a joint e↵ort of the universities of Oxford, Manchester and Edinburgh and around 20 industrial partners.

OXPath System: OXPath is a language designed for scalable web data extraction, crawling and automation. OXPath extends XPath with actions (e.g., click, form filling), Kleene star for iteration, and markers for data extraction. The Vadalog system allows for easy use of the OXPath system.

In the LAMBDA framework, **the Vadalog System will be tested in different scenarios with PUPIN data.** Staff exchange will be organized end of 2019 or beginning of 2020.

5. LAMBDA Big Data Analytics School Programme

5.1 LAMBDA Knowledge repository

The Work Package 3 (Open Education: Cooperation for Teacher and PhD Student Training) foresees establishment of a Knowledge repository in order to facilitate spreading learning materials, as well as exchange of best practice between research institutions from South-Eastern Europe and leading EU partners).
The new learning materials prepared by the internationally-leading counterparts will substantially improve the existing teaching courses and make them accessible via the SlideWiki.org platform.
The Knowledge repository is currently accessible for registered users of the LAMBDA platform only, who have been granted with 'associate partner' credentials. It has been divided into the following subsections.

https://project-lambda.org/Knowledge-repository/Lectures_private
https://project-lambda.org/Knowledge-repository/Datasets
https://project-lambda.org/Knowledge-repository/Tools
https://project-lambda.org/Knowledge-repository/Projects

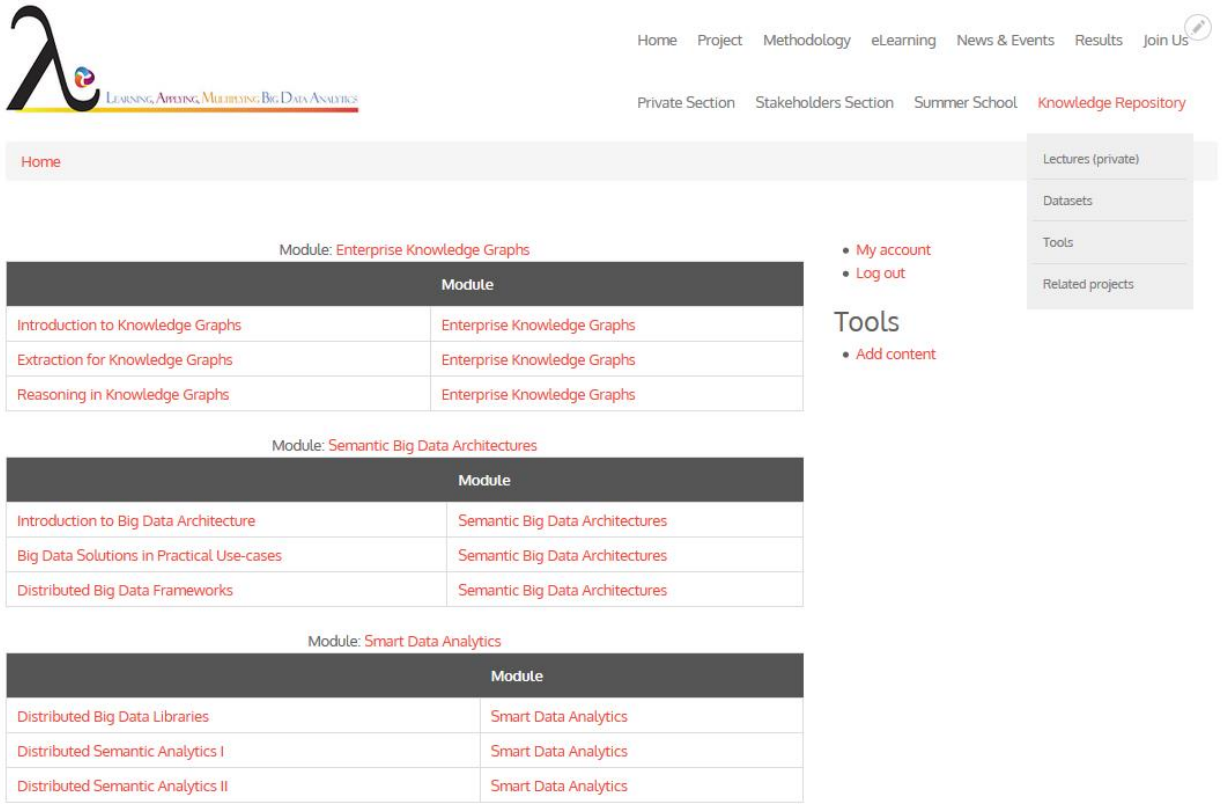


Figure 1. LAMBDA Lectures

5.2 Draft Agenda for the Big Data Analytics School 2019

This (3 days) event, scheduled for 18-20 June 2019[21], will be organized on an international level aiming at students, graduate students, as well as academic and industrial practitioners and researchers starting or being active in the field of Big Data. Within the draft program, topics have been selected that offer inspiring insights into the diverse field of Big Data. Invited teachers

[21] https://project-lambda.org/Announcement-1

(members of the LAMBDA Advisory Board) and LAMBDA teachers from UOXF, UBO and IAIS together will present lectures that provide an overview of relevant approaches and solutions and cover the basics, as well as advanced approaches to working with large and complex amounts of data. In 2019, the 3-day event will be scheduled as follows:

- 1st day LAMBDA Research-Industry Forum - Keynotes + Presentations from Companies
- 2nd day Big Data Analytics Summer School - Invited Lectures and Lectures from LAMBDA partners (UBO and UOXF)
- 3nd day Big Data Analytics Summer School - Lectures from LAMBDA partners (UBO and UOXF

Currently, the following experts have confirmed their keynotes / lectures:

- Sören Auer, TIB (Keynote) at the Research-Industry Forum, will participate at the Advisory Board Meeting
- Atanas Kiryakov, OntoText (Keynote) at the Research-Industry Forum, will participate at the Advisory Board Meeting
- Maria Esther Vidal, TIB (Invited Lecture) at-Big Data Analytics School
- Mari Carmen, UPM (Invited Lecture) at-Big Data Analytics School, , will participate at the Advisory Board Meeting

Additionally, commercial software providers from Serbia will be invited to present solutions. Stakeholders that confirmed their interest are SAS Serbia (SAS University Edition[22] and Cisco Big Data and Analytics solutions[23]).

LAMBDA consortium lectures will be splitted into 3 modules as follows.
Module 1 Enterprise Knowledge Graphs: Training materials will include formal conceptual frameworks for designing and maintaining knowledge graphs; such as strategies for the semiautomatic construction of such graphs from the combination of proprietary enterprise data and relevant public domain knowledge; opportunities and implications in terms of performance and access control.
Module 2 Semantic Big Data Architectures: Training materials will include approaches for better supporting the variety dimension of Big Data comprising RDF, RDF-Schema and OWL knowledge representation formalisms, mapping standards such as R2RML, JSON-LD and CSVW, the SPARQL query language etc. Integrating semantic and Big Data technologies can help making Big Data architectures and applications more flexible, adaptive and their implementation more efficient.
Module 3 Smart Data Analytics: Training materials will include different algorithms and tools related to Distributed Semantic Analytics, Semantic Question Answering, Structured Machine Learning, Deep Learning, Software Engineering for Data Science, Semantic Data Management, Knowledge Extraction and Validation.

Table 6. Draft Agenda of the Belgrade Big Data Analytics School 2019

1st Day (Research-Industry Forum)	2nd Day	3rd Day
Keynote - Research topics (TIB)	Invited Lecture (TIB)	Invited Lecture (UPM)
Big Data Research at the University of Oxford	LAMBDA Lectures	LAMBDA Lectures

[22] http://www.sas.com/en_us/software/university-edition.html
[23] UCS Common Platform Architecture for Big Data, http://www.cisco.com/c/en/us/solutions/data-center-virtualization/big-data/index.html

Big Data Research at the University of Oxford		
Invited Presentation		
Lunch Break		
Keynotes - Industry (OntoText)	LAMBDA Lectures	LAMBDA Lectures
Invited Presentation		
Invited Presentation		

5.3 Draft Agenda for the Big Data Analytics School 2020

The goal of the Research-Industry Forum is to build bridges with new stakeholders from academia, government and business and engage them in discussion of societal challenges relevant for Serbia, the region and Europe. The goal of the Foresight Exercise Workshop is to identify trends and drivers for the data-driven economy. The conclusions and recommendations from LAMBDA events will be public and available for decision and policy makers (e.g. for other science and technology Incubators in Serbia and abroad).

Table 7. Draft Agenda of the Belgrade Big Data Analytics School 2020

1st Day (**Foresight Exercise**)	2nd Day	3rd Day
Keynote - Research topics (representative of Big Data Value Association)	Invited Lecture (SZTAKI)	Invited Lecture (WUT)
Invited Presentation	LAMBDA Lectures	LAMBDA Lectures
Invited Presentation		
Invited Presentation		
Lunch Break		
Foresight Workshop	LAMBDA Lectures	LAMBDA Lectures

6 LAMBDA Staff Exchange

Starting from month six, in WP4 (Experts Exchange programme) framework, short-term visits of PUPIN staff (senior researchers and PhD students) to Fraunhofer, UBO and UOXF is planned. The activity foresees virtual brainstorming sessions and joint work of teams (virtual exchange of know-how and experience, mentoring activities, testing of existing tools, elaboration of ideas, and prototype design) as a first step towards strategic scientific and innovation partnerships.

In order to coordinate the work in WP2 framework, a staff exchange was organized also in M04, when the Coordinator Valentina Janev visited University of Bonn and Fraunhofer IAIS on 24 October 2018. The meeting was co-located with the 1st CognitiveEng-Net meeting for a project financed by BMBF.[24]

The Table below gives a draft plan for the main staff exchange until the end of the project.

Table 8. Draft Agenda of the Belgrade Big Data Analytics School 2020

When	Topics	Location	Duration
January / February 2019	Plenary Meeting 1 PUPIN PhD students visit IAIS - "Big Data tools and Industrial Analytics"	Bonn, Germany	5 days
May 2019	PUPIN PhD students visit UBO - "CSCUBS 2019 Conference"	Bonn, Germany	10 days
June 2019	IAIS, UBO, UOXF teachers visit PUPIN and give lectures during the BDA School 2019	Belgrade Serbia	5 days
January / February 2019	Plenary Meeting 3 PUPIN PhD students visit UOXF "VADA System"	Oxford, UK	10 days
June 2020	IAIS, UBO, UOXF teachers visit PUPIN and give lectures during the BDA School 2020	Belgrade Serbia	5 days

Additional events will be organized by UBO / IAIS staff in 2019 where PUPIN staff will take part in order to boost proof-of-concept development. As part of the BOOST 4.0 project a series of workshops will be organized as is presented in the Table below.

When	Topics	Location	Duration
11–12 April 2019	"Challenges of Trust, Interoperability and Autonomy in Industry 4.0"	TBD	2 days
8–12 July 2019	"Summer school / hackathon"	TBD	5 days

[24] https://project-lambda.org/CognitiveEng-net-1

2–4 September 2019	"Artificial Intelligence Solutions for Trustful Industry 4.0"	TBD	2 days

7 Involvements of Stakeholders at LAMBDA Events

One of the goals of the LAMBDA project is to spread excellence and disseminate knowledge throughout the West Balkan and South-East European countries, and to position PUPIN as a Big Data & Analytics HUB for the whole region. To that aim, networking events between PUPIN researchers and representatives of the Southeast Europe stakeholders (universities, research institutions, ICT industry and SMEs) will take place both in Belgrade and in other cities and towns of the region. Besides networking and dissemination activities, stakeholders has been invited as participant to the Big Data Analytics School. In order to further strengthen the links with stakeholders, the private part of the LAMBDA portal where information about events and training materials is stored is opened for registered users as well. LAMBDA project foresees adoption of LAMBDA outputs (training materials and lectures) by 10 Universities in the region by the end of the project including

- the School of Electrical Engineering, University of Belgrade, Serbia
- the Belgrade Metropolitan University, Serbia
- the Ss. Cyril and Methodius University, FYR Macedonia
- the University of Montenegro, Faculty of Electrical Engineering, Montenegro
- the University of Niš, Faculty of Mechanical Engineering, Serbia
- the University of Novi Sad, Faculty of Agriculture Serbia
- the University of Sarajevo, School of Economics and Business, Bosnia and Herzegovina
- the University of Tuzla, Faculty of natural sciences and mathematics, Bosnia and Herzegovina
- the University of Zagreb,Faculty of Organization and Informatics, Croatia

Figure 2 illustrates the Stakeholders database[25], users that currently have access to the private part of the LAMBDA portal.

[25] https://project-lambda.org/Stakeholders

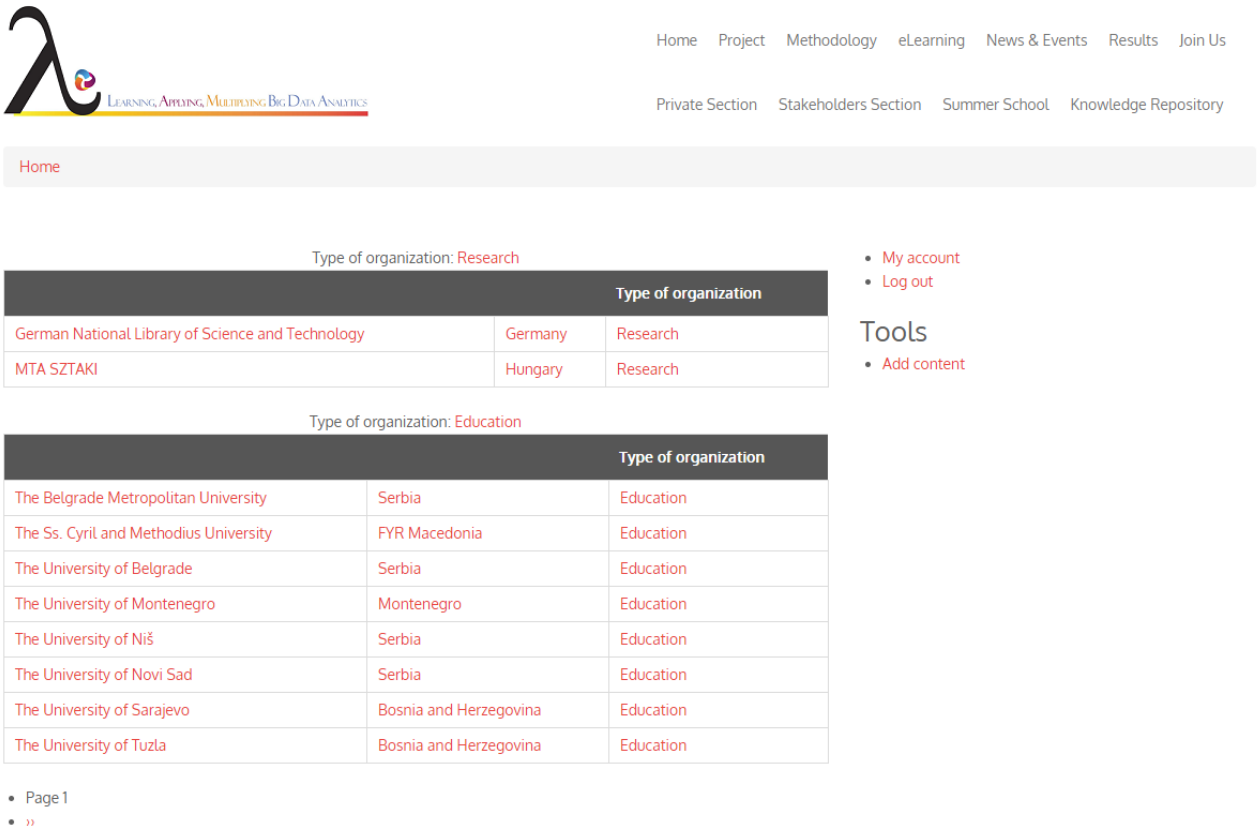


Figure 2. LAMBDA Stakeholders database

Additionally to assess the demand for education and training of these faculties, a survey has been initiated. Annex 1 points to data that is collected with the survey. The results of the survey will be analysed in period January - May 2019 when the final version of training materials is expected to be delivered.

Annex 1: Survey: Assessing the demand for Education and Training

Country Institute/Faculty name Description Research / Education / …	
Number of scientific employees Number of administrative employees Total number of employees	
Core competences of the Group and the experts who will be involved in LAMBDA Main R&D activities of the Group and the experts who will be involved in LAMBDA The larger three research projects (domain, duration and fund)	
Teaching activities/higher education (e.g. courses related to Data Analytics, AI, Big Data Architecture, Semantic Technologies, Big Data Analytics, Visualisation)	
The larger three Industrial projects (domain, and product) please provide links	
Relevant internal KPIs with the values for assessing the potential impact LAMBDA can make - Institute/Faculty level (eg. number of faculties on University level, number of enrolled/graduated students, number of publications…) Relevant external KPIs with the values for assessing the demand / shortage of IT experts / data analyst in your region/country	
Other Activities related to Big Data	
Communities and partnerships	
Strengths - Institute/Faculty level (e.g. link to Industry, number of professors,) Obstacles - Institute/Faculty level (e.g financial support, infrastructure)	