



LEARNING, APPLYING, MULTIPLYING BIG DATA ANALYTICS

Horizon 2020 Grant Agreement No 809965
Contract start date: July 1st 2018, Duration: 36 months

LAMBDA Deliverable 3.7

Belgrade BDA School (Sustainability Plan)

Due date of deliverable: 31/12/2020
Actual submission date: 30/12/2020

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, Marko Batić (PUPIN)
Contributor(s)	
Internal Reviewer(s)	Emanuel Sallinger (UOXF)
Approval Date	
Remarks	

Workpackage	WP 3 Cooperation for Teacher and Student Training
Responsible for WP	Institute for Computer Science - University of Bonn
Deliverable Lead	Institute Mihajlo Pupin (Valentina Janev)
Related Tasks	Task 3.3 Belgrade BDA School

Document History and Contributions

Version	Date	Author(s)	Description
0.1	01.12.2020	Valentina Janev	First draft
0.2	21.12.2020	Nikola Tomašević	Update
0.3	28.12.2020	Marko Batić	Update
0.4	28.12.2020	Emanuel Sallinger	Review
0.5	29.12.2020	Valentina Janev	Final version

© Copyright the LAMBDA Consortium. The LAMBDA Consortium comprises:

Institute Mihajlo Pupin (PUPIN)	Co-ordinator	Serbia
Fraunhofer Institute for Intelligent Analysis and Information Systems (Fraunhofer/IAIS)	Contractor	Germany
Institute for Computer Science - University of Bonn (UBO)	Contractor	Germany
Department of Computer Science - University of Oxford (UOXF)	Contractor	UK

Disclaimer:

The information in this document reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/her sole risk and liability.



Executive Summary

This report presents a plan for sustaining LAMBDA open education and training activities. One of the major knowledge transfer events of the LAMBDA project is the Belgrade Big Data Analytics Summer School, hence the activities necessary for organizing the next editions of the Belgrade Big Data Analytics Summer School are presented. Additionally, the efforts and investments needed for sustaining this event in the short and medium term (5-7 years) are presented.

In the LAMBDA project framework, the consortium partners (Institute Mihajlo Pupin, Fraunhofer Institute for Intelligent Analysis and Information System), Institute for Computer Science - University of Bonn and Department of Computer Science - University of Oxford) organized two editions of the LAMBDA Big Data Analytics Summer School. The target audience of the events were employees from the Institute Mihajlo Pupin interested in the adoption of emerging technologies in innovative industry scenarios, university staff interested in the adoption of lectures in master and PhD programs as well as professionals from industry and government sector.

The first edition of the Summer School ([Big Data Analytics Summer School, Belgrade, Serbia, June 2019](#)) provided introductory training in the field of Knowledge Graphs, Big Data Architectures, and Big Data Analytics. The topics were introduced by influential keynote speakers and professors.

The second edition of the Summer School ([Big Data Analytics Summer School, Belgrade, Serbia, June 2020](#)) provided advanced training in the field of Knowledge Graphs and Big Data processing. Lectures introduced methods for [Creation of Knowledge Graphs](#), [Data Lakes and Federated Query Processing](#), [Knowledge Graph Embeddings](#), [Context-Based Entity Matching for Big Data](#), [Scalable Knowledge Graph Processing](#), and others. Besides LAMBDA researchers speakers at the event were invited professors and guests from Europe and India.

To ensure further use of LAMBDA results after the completion of the project, the series of Summer School will be sustained by the Institute Mihajlo Pupin is a part of current EU research projects. To this aim, the 3rd edition of the school will be devoted to the energy sector with a focus on the use of semantic technologies and knowledge graphs for improving interoperability between stakeholders, the use of advanced Big Data algorithms and tools for improving the efficiency and accuracy of energy services and the impact of Big data on this domain.

LAMBDA partners (the University of Oxford, University of Bonn, Fraunhofer IAIS) and associated partners who actively contributed to the first two editions of the school (the German National Library of Science and Technology, OntoText, West University of Timisoara, Hungarian Academy of Sciences) are also committed to sustaining this activity. Hence, this document gives more details about the preparatory work underway related to the 3rd edition of the [Big Data Analytics Summer School, Belgrade, Serbia, June 2021](#) which focuses on energy.



Table of Contents

Executive Summary	3
Table of Contents	4
Abbreviations and Acronyms	5
List of Figures	5
List of Tables	5
1. Introduction	6
1.1 Objectives	6
1.2 Relation to Other Deliverables	7
1.3 Structure of the Deliverable	8
2. Background: Big Data Analytics Summer School	9
2.1 Year 2019: 1 st edition of LAMBDA Big Data Analytics Summer School	9
2.2 Year 2020: 2 nd edition of LAMBDA Big Data Analytics Summer School in 2020	10
2.3 Summary about participants and trainees	12
3. LAMBDA OERs: Current Status and Progress	13
3.2 Adoption of lectures for different industries	14
3.3 Searching the Lectures Repository	15
3.4 LAMBDA Book	16
3.5 Sustainability Plans of the LAMBDA Lectures	18
4. Future Plans and Timeframe	19
3.1 Big Data Analytics Summer School 2021	19
3.1.1 Year 2021 Organization	19
3.1.2 Year 2021 Topics	19
3.1.3 Draft Agenda	20
3.1.4 PhD Workshop	21
3.2 Big Data Analytics Summer School 2022-2025	21
5. Communication within the BDA-School Participants Group	22
5.1 Preparatory Activities	22
5.2 Main Communication Channels	22
6. Conclusion	23



Abbreviations and Acronyms

BDA	Big Data Analytics
CSVW	CSV on the Web
EKGs	Enterprise Knowledge Graphs
NoE	Network of experts
OERs	Open Educational Resources
RDF	Research Description Framework
R2RML	RDB to RDF Mapping Language
WP	Work Package

List of Figures

Figure 1. BDA School 2020 - Screenshot from the Opening session.....	7
Figure 2. Example of combining the LAMBDA Lectures in a form of a Curriculum	14
Figure 3. BDA School 2020 - Lectures	15
Figure 4. Example of Video Lecture	16
Figure 5. LAMBDA Book – Table of Contents	17
Figure 6. Sustainability analysis	21
Figure 7. Advisory Board Meeting, November 2020	22

List of Tables

Table 1. Target audience for this deliverable.....	8
Table 2. BDA School 2019 – Keynotes and Lectures	10
Table 3. BDA School 2020 – Keynotes and Lectures	11
Table 4. BDA School 2019 – Statistics about participants	12
Table 5. BDA School 2020 – Statistics about participants	12
Table 6. Overview of categories of lectures.....	13
Table 7. Example of elaboration of courses and curriculum	14
Table 8. LAMBDA Sustainability Plans regarding learning materials	18
Table 9. <i>Program of the School – Day 1, Day 2, Day 3</i>	20



1. Introduction

1.1 Objectives

The LAMBDA sustainability activities shall ensure the sustainability of the project services and the rapid and wide adoption of results beyond the project's lifetime and the EC funding. The primary goal of the sustainability plans, see also Deliverable 5.8 [Sustainability Measures and Activities, Innovation and IPR Management \(Plan\)](#), is:

- to further extend the use, implementation and development of the LAMBDA platform and services¹,
- to sustain the prototyping and development activities started in the Fraunhofer-PUPIN IAIS framework
- to propose and implement specific actions such as the [Big Data Analytics Summer School, Belgrade, Serbia, June 2021](#) which promote the exploitation of LAMBDA results.

This report presents a plan for sustaining LAMBDA open education and training activities. Taking into account that one of the major knowledge transfer event of the LAMBDA project is the Belgrade Big Data Analytics Summer School, this deliverable presents the activities necessary for organizing the next editions of the Belgrade Big Data Analytics Summer School. Additionally, the efforts and investments needed for sustaining this event in the short and medium term (5-7 years) are presented.

Overall, the sustainability of a bigger knowledge transfer event such as the LAMBDA Big Data Analytics Summer School that was organized in 2019 at Institute Mihajlo Pupin's premises (attended by over 60 participants from Serbia and Wes Balkan) requires an efficient and effective set of activities for

- planning the organization of the event;
- preparing teaching materials;
- attracting interesting lectures from outside the network;
- conducting the event and ensuring the necessary infrastructure (facilities, video-streaming, etc);
- boosting the motivation of the participating organisations to reuse the acquired knowledge by supporting the network with additional activities, e.g. webinars during the year.

Hence, the sustainability planning of the Big Data Analytics Summer School is performed at two parallel and complementary levels:

- Internally in the consortium organisations, i.e. through ensuring that all consortium participants will continue to use and expand the Lectures Repository following the end of the project;
- Externally through attracting and engaging third parties in the use of the Lectures Repository. Such an engagement is currently in PUPIN plans as part of the SINERGY project ([Capacity building in Smart and Innovative eENERGY management](#), GA No. 952140) that is expected to start in January 2021.

¹ <https://project-lambda.org/Knowledge-repository/Lectures>

1.2 Relation to Other Deliverables

This Deliverable is related to other WP3 reports including

1. [Deliverable 3.1 The 'Trainers' Network' Infrastructure²](#) that describes the 'Train-the-Trainers' approach adopted by LAMBDA.
2. [Deliverable 3.2 Enterprise Knowledge Graphs³](#) that summarizes the lectures from the 1st Big Data Analytics Summer School.
3. [Deliverable 3.3 Semantic Big Data Architecture⁴](#) that points to the lectures related to architectures.
4. [Deliverable 3.4 Smart Data Analytics⁵](#) that describes the complete set of lectures developed by June 2020.
5. [Deliverable 3.5 Belgrade BDA School \(Report 1\)⁶](#) that summarizes the organization of the 1st LAMBDA Big Data Analytics Summer School, <https://project-lambda.org/Summer-School-2019>, organized in Belgrade between June 17th and June 20th,
6. [Deliverable 3.6 Belgrade BDA School \(Report 2\)⁷](#) that summarizes the organization of the 2nd LAMBDA Big Data Analytics Summer School, organized online on 16th and 17th of June 2020, see <https://project-lambda.org/BDA-Summer-School-2020> (see Figure 1).

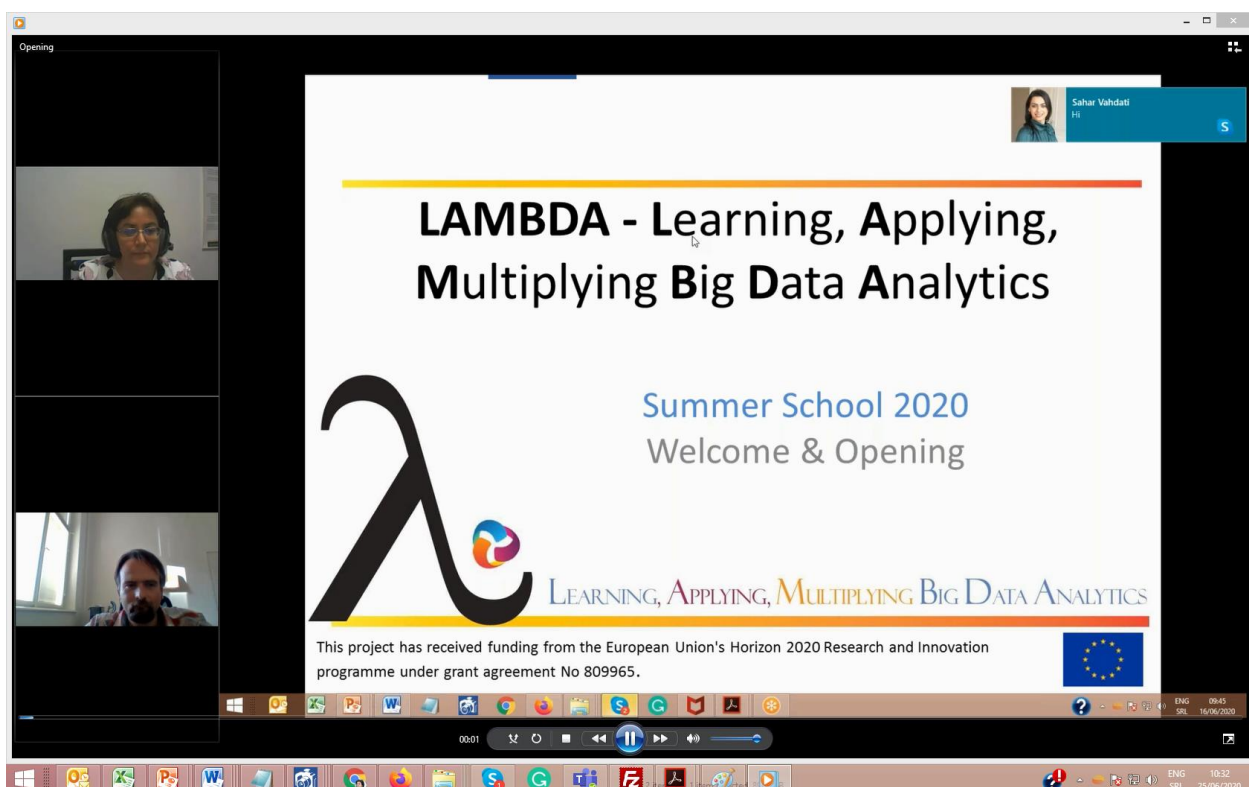


Figure 1. BDA School 2020 - Screenshot from the Opening session

² <https://project-lambda.org/D3.1>

³ <https://project-lambda.org/D3.2>

⁴ <https://project-lambda.org/D3.3>

⁵ <https://project-lambda.org/D3.4>

⁶ <https://project-lambda.org/D3.5>

⁷ <https://project-lambda.org/D3.6>

*Table 1. Target audience for this deliverable*

Intended audience	Reasons for interest in this document
Consortium partners	To summarize the conducted activities and use as a basis for the preparing the next year Summer School
Participants / Stakeholders	To have on one place the information about the event with links to all relevant support documents (open resource lectures and PPT presentations from the Advisory Board Members, links to the private side of the portal), and evidence (pictures, Agenda, Information Pack)
European Commission	To review and assess this deliverable as a required report based on the grant agreement
General public	To be informed about the LAMBDA activities

1.3 Structure of the Deliverable

This Deliverable is organized as follows

- Background and statistics about previous editions of the Big Data Analytics Summer School (Section 2);
- Current status of LAMBDA Repository of Lectures (Section 3);
- Tentative Agenda for the [Big Data Analytics Summer School, Belgrade, Serbia, June 2021](#) with list of teachers who will be invited to collaborate in the following years (Section 4);
- Communication channels (Section 5).



2. Background: Big Data Analytics Summer School

2.1 Year 2019: 1st edition of LAMBDA Big Data Analytics Summer School

The 1st edition of LAMBDA Big Data Analytics Summer School, <https://project-lambda.org/Summer-School-2019>, was organized in Belgrade between June 17th and June 20th, 2019, see Agenda⁸. Overall, more than 60 participants gathered at the PUPIN premises to exchange knowledge and expertise in Big Data technologies. The objective of the summer school was to give the PhD students and experts from Serbia and abroad (see also LinkedIn NoE)⁹ and PUPIN researchers an opportunity to learn about the newest technologies and trends in this and related fields from respectable professors, as well as to hear about use cases from influential tech companies such as OntoText, SAS Institute, CISCO, Meltwater, and DeepReason.ai. The program contained presentations by well-known international experts, members of the LAMBDA Advisory Board and the LAMBDA consortium. The topics of the keynotes and the lectures covered [Enterprise Knowledge Graphs](#), [Semantic Big Data Architectures](#), [Smart Data Analytics](#), and Big Data use cases from different domains.

Organizers:

- Valentina Janev (PUPIN)
- Damien Graux (Fraunhofer IAIS)
- Hajira Jabeen (UBO)
- Emanuel Sallinger (UOXF)

Keynotes:

- [Sören Auer, Director, German National Library for Science and Technology](#) (Germany)
- [Atanas Kiryakov, CEO, OntoText](#) (Bulgaria)
- [Maria Esther Vidal, Head of Scientific Data Management Research Group, German National Library for Science and Technology](#) (Germany)

Invited Speakers:

- Daniel Pop, [Research Institute e-Austria Timisoara](#) / [West University of Timisoara](#)
- Gabriel Iuhasz, [West University of Timisoara](#)
- [Radenko Čitaković, CISCO - Serbia](#)
- [Nikola Nikačević, Analytics advisor, SAS Institute doo, Serbia](#)
- Luigi Bellomarini, Banca d'Italia and [University of Oxford](#)
- Tim Furche, Meltwater Inc and [University of Oxford](#)

Speakers from core group (UBO, UOXF, Fraunhofer, PUPIN):

- Emanuel Sallinger, [University of Oxford](#) and TU Wien
- [Damien Graux, Fraunhofer Institute for Intelligent Analysis and Information Systems, Germany](#)
- Hajira Jabeen, University of Bonn
- [Dea Pujić, Institute Mihajlo Pupin](#)
- [Marko Jelić, Institute Mihajlo Pupin](#)

⁸ https://project-lambda.org/sites/default/files/2019-06/BDA_2019_Information_Pack_and_Agenda.pdf

⁹ <https://www.linkedin.com/groups/12129621/>



Table 2. BDA School 2019 – Keynotes and Lectures

Module	Lecture	Contributed by
Case Studies	Open Research Knowledge Graph	TIB (keynote)
Big Data and KGs Tools	GraphDB: Use Cases, Analytics and Linking	OntoText (keynote)
Case Studies	Precision Medicine- A Use Case	TIB (keynote)
Enterprise Knowledge Graphs	Introduction to Knowledge Graphs	UOXF
Enterprise Knowledge Graphs	Extraction for Knowledge Graphs	UOXF
Semantic Big Data Architectures	Introduction to Big Data Architecture	Fraunhofer , UBO
Semantic Big Data Architectures	Big Data Solutions in Practical Use-cases	Fraunhofer , UBO
Semantic Big Data Architectures	Distributed Big Data Frameworks	UBO , Fraunhofer
Smart Data Analytics	Distributed Big Data Libraries	UBO , Fraunhofer
Smart Data Analytics	Distributed Semantic Analytics I	UBO , Fraunhofer
Smart Data Analytics	Distributed Semantic Analytics II	UBO , Fraunhofer
Smart Data Analytics	SANSA - Scalable Semantic Analytics Stack	UBO
Case Studies	Data Analytics for Energy Sector	PUPIN
Big Data and KGs Tools	Data Science with Spark and Hadoop	UBO
Big Data and KGs Tools	Spark using Scala	UBO
Foundations	Big Data Outlook, Tools, and Architectures	UBO

2.2 Year 2020: 2nd edition of LAMBDA Big Data Analytics Summer School in 2020

The 2nd edition of LAMBDA Big Data Analytics Summer School was organized on June 16th and June 17th, 2020, see Agenda¹⁰. Because of COVID-19, the event took place online. Overall, more than 70 participants joined the online sessions. The topics of the keynotes and the lectures extended the topics introduced in 2020 ([Enterprise Knowledge Graphs](#), [Semantic Big Data Architectures](#), [Smart Data Analytics](#)) and included also [Foundations](#), [Artificial intelligence](#), [Big Data Tools](#) and [Case Studies](#). The website of the summer school, <https://project-lambda.org/BDA-Summer-School-2020>, provides more details about the organization and topics discussed at the school.

Organizers:

- Valentina Janev (PUPIN)
- Damien Graux (Trinity College Dublin and Fraunhofer IAIS)
- Hajira Jabeen (UBO)
- Emanuel Sallinger (UOXF)

Keynotes:

- [Georgios Paliouras, NCSR “Demokritos”](#) (Greece)
- [Mariana Damova, Mozaika](#) (Bulgaria)
- [Gloria Bordogna, Italian National Research Council IREA](#) (Italy)

¹⁰ https://project-lambda.org/sites/default/files/2020-06/BDA_2020_Information_Pack_and_Agenda_3.pdf

**Invited Speakers:**

- [Maria Esther Vidal, German National Library for Science and Technology](#)
- [Anastasia Dimou, Ghent University](#)
- [Diego Collarana, Fraunhofer IAIS](#)
- [Hajira Jabeen, University of Bonn](#)
- [Debasis Das, Indian Institute of Technology](#)
- [Sahar Vahdati, University of Oxford](#)

Speakers from core group (UBO, UOXF, Fraunhofer, PUPIN):

- [Emanuel Sallinger \(Plenary Talk\), University of Oxford](#)
- [Valentina Janev, Institute Mihajlo Pupin](#)
- [Damien Graux, Trinity College Dublin and Fraunhofer IAIS](#)

Technical support:

- [Dea Pujić, Institute Mihajlo Pupin](#)
- [Marko Jelić, Institute Mihajlo Pupin](#)

Table 3. BDA School 2020 – Keynotes and Lectures

Module	Lecture	Contributed by
Case Studies	Semantic information infrastructures from business information delivery to water management	Mozaika
Case Studies	Soft computing for Transparent synthesis of Geo Big Data	IREA-CNR
Enterprise Knowledge Graphs	Knowledge Graph Embeddings	UOXF
Enterprise Knowledge Graphs	Creation of Knowledge Graphs	UGENT
Enterprise Knowledge Graphs	Reasoning in Knowledge Graphs: An Embeddings Spotlight	UOXF
Semantic Big Data Architectures	Data Lakes and Federated Query Processing	TIB
Smart Data Analytics	Scalable Knowledge Graph Processing using SANSA	UBO, Fraunhofer
Case Studies	Chronorobotics - Spatio-temporal models for social and service robots	CVUT-CZ
Case Studies	IntelliSys: Intelligent System for Road Safety	IIT-IN
Case Studies	Reasoning on Financial Knowledge Graphs: The Case of Company Networks	UOXF
Artificial intelligence	Data for AI: Foresight	Fraunhofer
Artificial intelligence	AI and Knowledge Graphs	UOXF
Big Data and KGs Tools	Context-Based Entity Matching for Big Data	Fraunhofer
Survey	Survey on Big Data Applications	PUPIN



2.3 Summary about participants and trainees

Table 4. BDA School 2019 – Statistics about participants

Teachers / Speakers outside the consortium		Participants from West Balkan Countries	
Country	Number	Country	Number
Germany	2	Croatia	3
Bulgaria	1	Bosna and Herzegovina	2
Hungary	1	Montenegro	2
Romania	2	North Macedonia	2
Serbia	2		

Teachers / Speakers from the consortium		Participants from Serbia	
Country	Number	Organization	Number
Germany	3	PUPIN	22
UK / Germany	1	University of Niš	2
UK / Austria	1	University of Novi Sad	2
Italy	1	University Metropolitan	3
Serbia	2	University UNION	2
		Other Faculties from the University of Belgrade	3
		Other stakeholders (Banks, Government agencies, etc.)	5
Total	16	Total	48

Table 5. BDA School 2020 – Statistics about participants

Teachers and Participants			
Country	Number	Country	Number
Serbia	35	India	1
Italy	7	Greece	1
Germany	7	Bulgaria	1
UK	4	Romania	1
Czech Republic	2	North Macedonia	1
Belgium	2	Montenegro	1
Austria	2	Albania	1
Croatia	2	Ireland	1
Bosnia and Herzegovina	2		
Total	71		

Approximately one third of the trainees were from the Mihajlo PUPIN Institute (PhD students, senior researchers and engineers/professionals).



3. LAMBDA OERs: Current Status and Progress

In the course of the project more than 30 lectures have been developed that are available at

<https://project-lambda.org/Knowledge-repository/Lectures>

In the first project year 10 lectures were developed categorised as following:

1. [Enterprise Knowledge Graphs](#) (see **Deliverable 3.2**): lectures that include formal conceptual frameworks for designing and maintaining knowledge graphs; such as strategies for the semi-automatic 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.
2. [Semantic BD Architectures](#) (see **Deliverable 3.3**): lectures that 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 to make Big Data architectures and applications more flexible, adaptive and their implementation more efficient.
3. [Smart Data Analytics](#) (**Deliverable 3.4**): lectures that 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. Overview of categories of lectures

1 st project year	2 nd project year
Enterprise Knowledge Graphs ¹¹ Semantic Big Data Architectures ¹² Smart Data Analytics ¹³	Artificial Intelligence ¹⁴ Surveys ¹⁵ Foundations ¹⁶ Enterprise Knowledge Graphs Semantic Big Data Architectures Big Data and Knowledge Graphs Tools ¹⁷ Smart Data Analytics Case Studies ¹⁸

In the second project year additional 20 lectures were developed lectures and new categories of lectures were specified as is presented in Table 6. These lecture materials constitute one of the main valuable outcomes of the LAMBDA project.

¹¹ <https://project-lambda.org/Knowledge-Graphs>

¹² <https://project-lambda.org/Big-Data-Architectures>

¹³ <https://project-lambda.org/Smart-Data-Analytics>

¹⁴ <https://project-lambda.org/Artificial-Intelligence>

¹⁵ <https://project-lambda.org/Surveys>

¹⁶ <https://project-lambda.org/Foundations>

¹⁷ <https://project-lambda.org/Knowledge-Graphs-Tools>

¹⁸ <https://project-lambda.org/Case-Studies>

3.2 Adoption of lectures for different industries

Based on the customization of the LAMBDA learning material, different training and workshops could be carried out. Table 7 gives an overview of the adoption of LAMBDA lectures for different industries and courses, while Figure 2 presents an example of adoption of lectures for financial industry.

Table 7. Example of elaboration of courses and curriculum

Partner	Course
PUPIN	Knowledge Graphs and Big Data for Energy sector Knowledge Graphs and Big Data for eGovernment Business Intelligence course Semantic Web course Artificial Intelligence course
UOXF	Finance industry

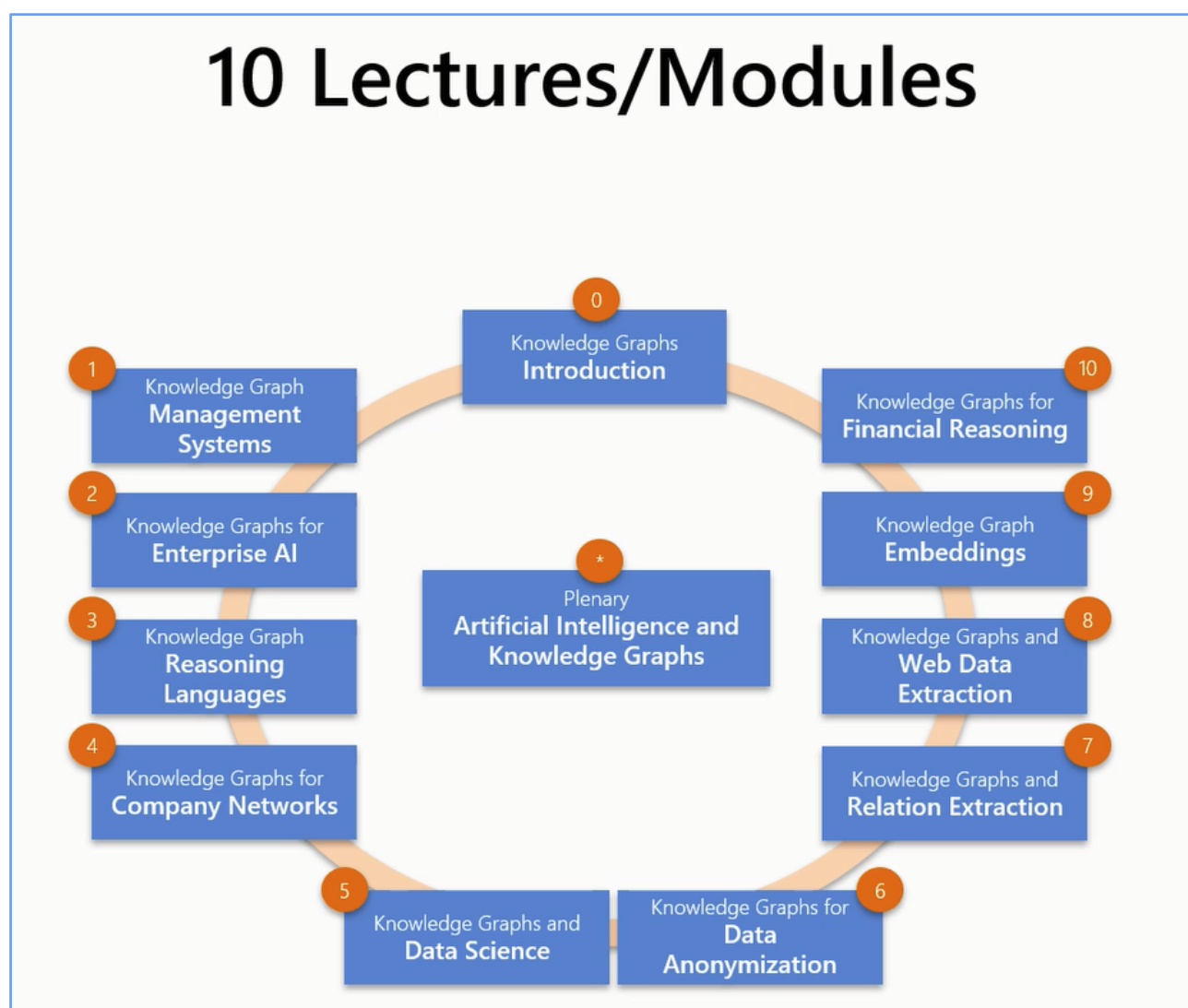


Figure 2. Example of combining the LAMBDA Lectures in a form of a Curriculum



3.3 Searching the Lectures Repository

The easiest way to retrieve the lectures is to use the Search functionalities of the LAMBDA Platform under this link <https://project-lambda.org/Knowledge-repository/Lectures>

The user has two options:

- Search by topic
- Search by event (select BDA School 2020), as is presented in Figure 3.

The screenshot shows the LAMBDA platform interface. At the top left is the LAMBDA logo with the tagline 'LEARNING, APPLYING, MULTIPLYING BIG DATA ANALYTICS'. To the right are navigation links: Home, Project, Summer School, eLearning, Private Section, Stakeholders Section, and Knowledge. Below the logo is a 'Home' button. The main search area has two dropdown menus: 'Select MODULE' (set to '- Any -') and 'Select Event' (set to '- Any -'). An 'Apply' button is next to the 'Select Event' dropdown. The 'Select Event' dropdown is open, showing options: '- Any -', 'BDA School 2019', 'BDA School 2020' (highlighted), and 'Other'. To the right of the search area are links for 'My account' and 'Log out', and a 'Tools' section with a link for 'Add content'. Below the search area is a table with the following data:

Module		Presented at event	Contributed by	Available as
Artificial intelligence	Data for AI: Foresight	BDA School 2020	Fraunhofer	Video, PPT
Case Studies	Semantic information infrastructures from business information delivery to water management	BDA School 2020		Video

Figure 3. BDA School 2020 - Lectures

Video lectures have also been uploaded to the LAMBDA YouTube Channel, <https://www.youtube.com/channel/UC9BCAGX1dzCI2akuRxILq6Q/> and are embedded in pages on the LAMBDA platform, see an example of embedding a video lecture in LAMBDA portal on Figure 4.

[View](#) [Edit](#) [Delete](#) [Revisions](#)Posted on: Wed, 04/01/2020 - 21:11 By: [valentina.janev](#)

The goal of this chapter is to shed light on different types of big data applications needed in various industries including healthcare, transportation, energy, banking and insurance, digital media and e-commerce, environment, safety and security, telecommunications, and manufacturing. In response to the problems of analyzing large-scale data, different tools, techniques, and technologies have been developed and are available for experimentation. In our analysis, we focused on literature (review articles) accessible via the Elsevier ScienceDirect service and the Springer Link service from more recent years, mainly from the last two decades. For the selected industries, this lecture also discusses challenges that can be addressed and overcome using the semantic processing approaches and knowledge reasoning approaches discussed in this book.

The Lecture has been presented at the Big Data Analytics Summer School 2020 by [Dr. Valentina Janev](#), [Institute Mihajlo Pupin](#), LAMBDA project coordinator.

- [My account](#)
- [Log out](#)

Tools

- [Add content](#)

P BDA Summer School 2020: Survey of Big Data Applications

Big Data Ecosystem

- The term **Ecosystem** is defined in scientific literature as a complex network or interconnected systems.
- While in the past corporations used to deal with static, centrally stored data collected from various sources, with the birth of cloud services, cloud computing is rapidly overtaking the traditional in-house system as a reliable, scalable and cost-effective IT solution.
- Thus, large datasets – log files, social media sentiments, click-streams – are no longer expected to reside within a central server or within a fixed place in the cloud.
- To handle the copious amounts of data, **advanced analytical tools are needed which can process and store**

Figure 4. Example of Video Lecture

3.4 LAMBDA Book

The LAMBDA consortium prepared a book that includes the lectures presented by the LAMBDA researchers at the 1st and the 2nd BDA School. The table of contents is presented on Figure 5.



Table of Contents

Foundations

Chapter 1 Ecosystem of Big Data	1
<i>Valentina Janev</i>	
Chapter 2 Knowledge Graphs: The Layered Perspective	19
<i>Luigi Bellomarini, Emanuel Sallinger, and Sahar Vahdati</i>	
Chapter 3 Big Data Outlook, Tools, and Architectures	34
<i>Hajira Jabeen</i>	

Architecture

Chapter 4 Creation of Knowledge Graphs	55
<i>Anastasia Dimou</i>	
Chapter 5 Federated Query Processing	69
<i>Kemele M. Endris, Maria-Esther Vidal, and Damien Graux</i>	
Chapter 6 Reasoning in Knowledge Graphs: An Embeddings Spotlight ...	82
<i>Luigi Bellomarini, Emanuel Sallinger, and Sahar Vahdati</i>	

Methods and Solutions

Chapter 7 Scalable Knowledge Graph Processing using SANSa	97
<i>Hajira Jabeen, Damien Graux, and Gezim Sejdiu</i>	
Chapter 8 Context-Based Entity Matching for Big Data	114
<i>Mayesha Tasnim, Diego Collarana, Damien Graux, and Maria-Esther Vidal</i>	

Applications

Chapter 9 Survey on Big Data Applications	139
<i>Valentina Janev, Dea Pujić, Marko Jelić, and Maria-Esther Vidal</i>	
Chapter 10 Case Study from the Energy Domain	155
<i>Dea Pujić, Marko Jelić, Nikola Tomašević, and Marko Batić</i>	
Bibliography	171

Figure 5. LAMBDA Book – Table of Contents



3.5 Sustainability Plans of the LAMBDA Lectures

The individual Sustainability Plans of the LAMBDA partners are summarized in Table 8.

Table 8. LAMBDA Sustainability Plans regarding learning materials

Partner	Lectures	Sustainability Plan
PUPIN	Big Data Ecosystem Survey on Big Data Tools Overview and Comparison of Machine Learning Algorithms Survey on Big Data Applications Open and Big Data – Utilization Perspective Data Analytics for Energy Sector Predictive Analytics in Renewable Energy Systems	PUPIN will sustain the activities and will prepare additional lectures relevant for <ul style="list-style-type: none">- Energy- Transport and Traffic- eGovernment- Industry 4.0 and Manufacturing
Fraunhofer	Data for AI: Foresight What is Knowledge Graph ? Data Lakes and Federated Query Processing Context-Based Entity Matching for Big Data Conversational AI The Revolution of AI Introduction to Big Data Architecture Big Data Solutions in Practical Use-cases	IAIS will continue expanding the lectures showing new use-case we work at Fraunhofer IAIS. Moreover, we will add additional material in the following two topics: <ol style="list-style-type: none">1. Conversational Agents from Big Data Sources2. Question Answering over Big Knowledge Graphs
UBO	Big Data Outlook, Tools, and Architectures SANSA - Scalable Semantic Analytics Stack Data Science with Spark and Hadoop Spark using Scala Distributed Big Data Frameworks Distributed Big Data Libraries Distributed Semantic Analytics I Distributed Semantic Analytics II Scalable Knowledge Graph Processing using SANSA	UBO will prepare additional relevant learning materials for: <ul style="list-style-type: none">- Distributed Machine Learning algorithms.- Scalable Clustering algorithms using SANSA.
UOXF	Knowledge Graphs: Introduction Knowledge Graph Management Systems Knowledge Graphs and Enterprise AI Knowledge Graph Reasoning Languages Knowledge Graphs for Company Networks Knowledge Graphs and Data Science Knowledge Graphs and Data Anonymization Knowledge Graphs and Relation Extraction Knowledge Graphs and Web Data Extraction Knowledge Graph Embeddings Knowledge Graphs for Financial Reasoning Plenary: Artificial Intelligence and KGs The Vadalog System	UOXF will, in collaboration with TU Wien and the Central Bank of Italy, develop additional learning material: <ul style="list-style-type: none">• extending existing lectures with additional content and material• developing additional learning material for KGs supporting a diversity of learners• providing a full course on “Knowledge Graphs”



4. Future Plans and Timeframe

This section discusses different options for sustaining the event on an annual basis. It proposes a tentative plan for the event in 2021, and a list of speakers for the forthcoming editions of the Summer School.

In year 2021, the group will start a new project [SINERGY – Capacity building in Smart and Innovative eENERGY management](#). In order to ensure a smooth transition of the activities from LAMBDA to SINERGY, we decided to focus the year 2021 school to energy topics.

3.1 Big Data Analytics Summer School 2021

3.1.1 Year 2021 Organization

Organizers:

- Valentina Janev (PUPIN)
- Diego Collarana Vargas (Fraunhofer IAIS)
- Jens Lehmann (UBO)
- Emanuel Sallinger (UOXF)

Keynotes:

- Philippe Calvez, ENGIE (France)

Invited Speakers:

- Maria Esther Vidal, Head of Scientific Data Management Research Group, German National Library for Science and Technology (Germany)
- Dr. Marcus Martin Keane, NUI Galway National University of Ireland, Galway (Ireland)
- Dr. Johannes Stöckl, AIT Austrian Institute of Technology (Austria)
- Erik Maqueda Moro, Tecnalia (Spain)
- Martino Maggio, Engineering Ingegneria Informatica spa (Italy)
- Sarra Ben-Abbes, ENGIE
- Lynda Themal, ENGIE
- Carsten Draschner, UBO
- Andrej Čampa, ComSensus
- Dr. Kemele Endris, German National Library for Science and Technology (Germany)
- Hantong Liu, Fraunhofer IAIS

Speakers from core group (UBO, UOXF, Fraunhofer, PUPIN):

- Emanuel Sallinger, [University of Oxford](#) and TU Wien
- Valentina Janev (PUPIN)
- Nikola Tomašević (PUPIN)
- Marko Batić (PUPIN)

3.1.2 Year 2021 Topics

The topics and program of the School has been decided by the core LAMBDA team based on the needs of PUPIN employees. Topics have been selected from the energy domain as follows:



- Big Data & Secure, Clean and Efficient Energy
 - Challenges in Energy Management Solutions (Production, Transmission and Consumption)
 - Renewable Energy Sources and Smart Energy Management
 - Smart Grids Management
 - Building Energy Management System
- Big Data & Analytical Services
 - Forecasting Models
 - Predictive Maintenance
 - Energy Usage Optimization
- Semantic Technologies for Energy
 - Semantic Models and Standards
 - Knowledge Graphs in the Energy Domain
 - Interoperability and European Data Spaces

3.1.3 Draft Agenda

The Program of the School is divided into 7 sessions named

1. Keynotes Session (Chair: Valentina Janev)
2. EU Perspective and Project Networking Event (Chair: Valentina Janev)
3. Foresight panel on BigData (Chair: Jens Lehmann)
4. Big Data Architectures and Knowledge Graphs for Energy Data Space (Chair: Diego Collarana)
5. Foresight panel on Energy (Chair: Nikola Tomašević)
6. Smart and Innovative eENERGY (Chair: Nikola Tomašević)
7. PhD Workshop (Chair: Marko Batić)
8. New Business Models – PUPIN Testbed (Chair: Marko Batić)
9. Closing Session (Chair: Valentina Janev)

as is presented in Table 9.

Table 9. Program of the School – Day 1, Day 2, Day 3

	Day 1	Day 2	Day 2
09:00	Registration / Establishing connections		
09:30	Opening	Opening	Opening
10:00	Session 1	Session 4	Session 7
	Keynotes	PLATOON lectures	PhD Workshop
	Lunch Break		
13:00	Session 2	Session 5	Session 8
	EU Perspective	Foresight Panel - Energy	PUPIN Testbed
15:00	Session 3 - Panel	Session 6	Session 9
	Foresight Panel – Big Data	SINERGY Lectures	Closing

3.1.4 PhD Workshop

In year 2021, in order to promote the work of young researchers, a PhD workshop¹⁹ will be organized, as part of the Big Data Analytics Summer School.

The following PhD students have confirmed their interest in submitting their works for this workshop:

- Dea Pujić
- Marko Jelić
- Marija Popović
- Katerina Stanković
- Dušan Popadić

3.2 Big Data Analytics Summer School 2022-2025

The focus of the first two editions of the BDA School was more on knowledge-driven technologies for different sectors. Starting from year 2021, we would like to provide more focused lectures, as is the case in year 2021 for the energy domain.

The organizations of the forthcoming editions of the Big Data Analytics Summer School in 2021, 2023, 2024 is ensured with the current resources of the Fraunhofer-PUPIN JPO²⁰, see running EU projects from the H2020 programme in Figure 6.

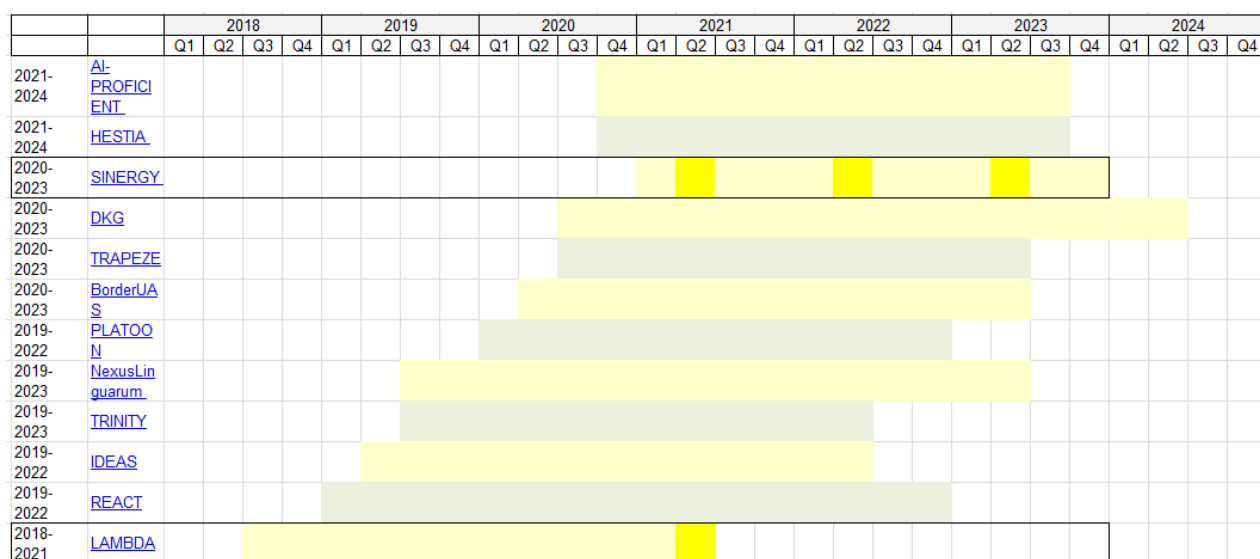


Figure 6. Sustainability analysis

¹⁹ <https://project-lambda.org/PhD-Workshop-2021>

²⁰ <http://www.pupin.rs/en/imp-organization/fraunhofer-pupin-jpo/>



5. Communication within the BDA-School Participants Group

5.1 Preparatory Activities

The preparation of BDA School 2021 started in November 2020.

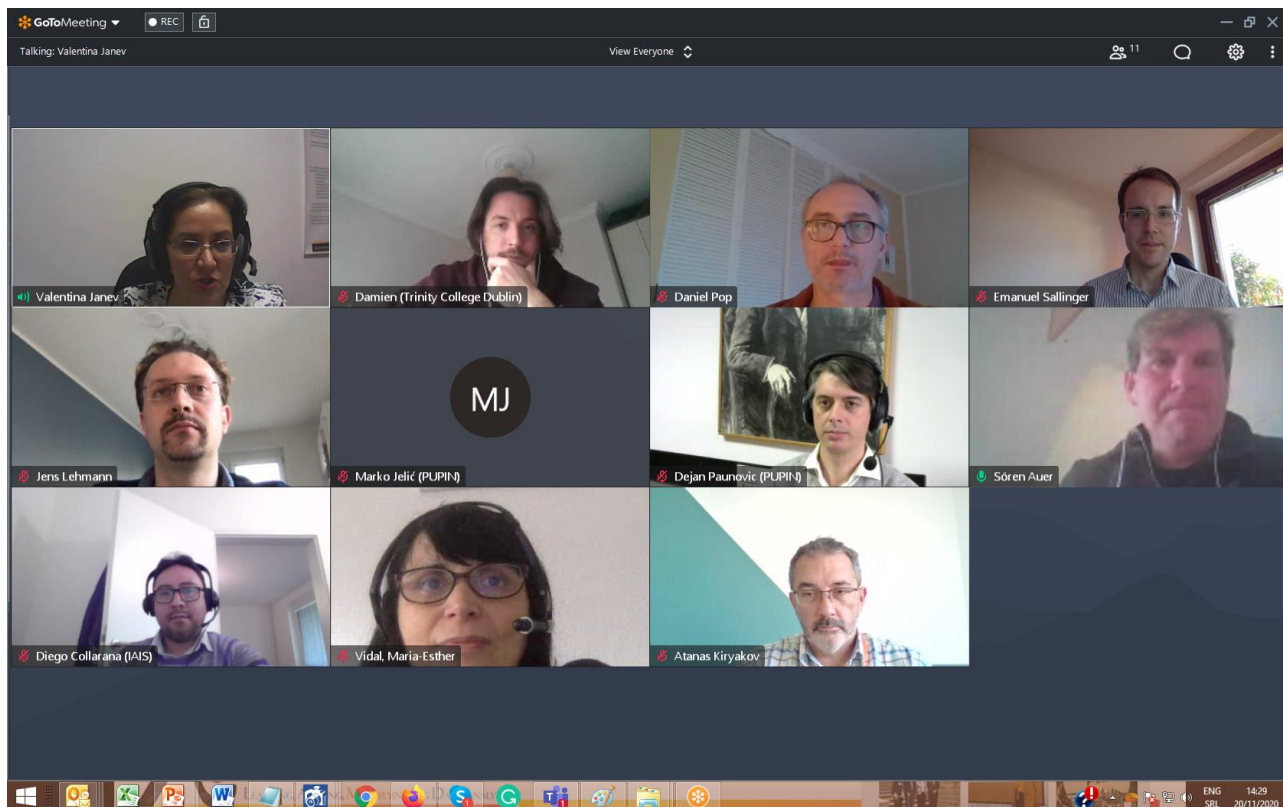


Figure 7. Advisory Board Meeting, November 2020

5.2 Main Communication Channels

- The LAMBDA Platform, see <https://project-lambda.org/Summer-Schools>
- The bda-school@mail.project-lambda.org as one of the main channels for information exchange
- The LinkedIn LAMBDA Network of Experts Group, <https://www.linkedin.com/groups/12129621/>
- The Twitter Account, <https://twitter.com/Net4LAMBDA>
- the Facebook Group, <https://www.facebook.com/valentina.janev.16>



6. Conclusion

This Deliverable gives an overview of sustainability activities related to the Big Data Analytics Summer School. More details about other sustainability activities have been given in Deliverable 5.8 [Sustainability Measures and Activities, Innovation and IPR Management \(Plan\)](#)²¹.

Overall, the experience gained from organizing the 1st and 2nd Summer School in Big Data Analytics is very positive. We see not only a clear benefit for the PUPIN employees but also for other participants at the School with heterogeneous educational background and professional level. The preparations for next year's summer school have already started and we anticipate the 3rd Summer School will take place in June 2021.

²¹ <https://project-lambda.org/D5.8>