Preface of the First Ph.D. Workshop on Big Data Analytics from the LAMBDA Network

Damien Graux[1] and Valentina Janev[2]

[1] Inria, Université Côte d'Azur, CNRS, I3S, France damien.graux@inria.fr
[2] Institut Mihajlo Pupin, University of Belgrade, Serbia
valentina.janev@institutepupin.com

Abstract. The Doctoral Workshop was organized with the aim to provide Ph.D. candidates with an opportunity to present their research projects and their scientific questions to their peers and international academic experts in the domain of Big Data analytics. Another goal of the Workshop was to develop a supportive community within which students can begin to develop their professional networks, interacting with peers and senior scholars from the field.
This 1st edition, co-located with the LAMBDA Big Data Analytics summer school located in Belgrade (Serbia), gathered the community around nine research publications, three talks and one invited keynote presentation. The event took place online on the 17th of June, 2021.

Keywords: Big Data · Analytics · Energy · Ph.D. Workshop · LAMBDA

The European H2020 LAMBDA Project

Big Data refers to data sets which have large size and complex structure. The data size can range from dozens of terabytes to a few zettabytes and is still growing [16]. Big Data Analytics, hence, refers to the strategy of analysing large volumes of data that are gathered from a wide variety of sources, including social networks, transaction records, videos, digital images and different kind of sensors. In an attempt to support the European data economy policy [10], our consortium proposed a training approach [13] and established the infrastructure for collaborative work of teachers/trainers with Ph.D. students and other interested parties such as industries. In particular, our activities are taking place within the scope of the Horizon 2020 project named LAMBDA, standing for "Learning, Applying, Multiplying Big Data Analytics".

Our overall objectives are to stimulate scientific excellence and innovation capacity, to increase the research capacities and to unlock the research potential in the ICT area in the whole West Balkan region, turning the partner-institutes into regional points of reference when it comes to multidisciplinary ICT competence related to Big Data analytics. In early 2018, the consortium started activities for improving the skills and competences for smart data management

Copyright © 2021 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

through a set of actions such as: the organization of international events (training, workshops, webinars, conferences), the development of a specific learning kit about Big Data Analytics [13], or the publication of an open-access book[1] [14].

In particular, to increase the audience and the availability of our initiative, we developed an online knowledge repository to help students acquiring new skills:

<https://project-lambda.org/Learning>

The learning materials that were produced are free, stored in a public repository and available online [12] using the OpenCourseWare platform SlideWiki [15]. This project repository aims at facilitating the exchange of learning materials, tools, project results and best practice between the international leading organizations and research institutions and industry from the West Balkan countries.

Since 2019, among the various activities of the project, the LAMBDA consortium organized three editions of the Big Data Analytics Summer School in Belgrade (Serbia) where students were able to connect with experts and to receive lectures and advice from them. This year, during the 2021 edition of the summer school, we organized for the first time a Ph.D. day as the community is getting momentum, in order to give Ph.D. students a place to present their current state of research efforts.

Scientific programme

The Ph.D. workshop started with the keynote entitled "Hints to Save Time when dealing with Big Data" given by Dr. Damien Graux from Inria[2] (France). He described how to adopt strategies which should help Big Data practitioners to fasten their processes and tasks [11]. In particular, Dr. Graux gave an overview of techniques and methods to better define the scope and type of Big Data the practitioners are dealing with. He gave indicators and rules to quickly know whether some systems should be considered or not for specific use cases. Overall, this keynote gave the audience in-depth details on practical use cases backed by cutting-edge research techniques.

Then, a session on Semantic Web tools in the context of Big Data happened. It was composed of two articles [4, 1]. Next a session dealing with big textual resources was organized. Three papers were presented: Gjorgjevikj *et al.* on detecting sustainable development indicators through text; Mishev *et al.* focused on text-to-speech for Macedonian; and [6] proposed a blockchain-based platform for logs of citizens' consents. After the lunch break, the third and fourth sessions gathered articles addressing energy challenges. Numerical tools for combustion behaviors predict were presented in [7]. In [8] and [2], authors respectively discussed how to detect faults in distribution networks and how to detect events in power cables. A machine learning based wind turbine production forecaster

[1] Downloaded more than 65 000 times as of December 2021.

[2] https://dgraux.github.io/

was then presented by Pujić and Janev. Efforts to develop a control platform dedicated to renewable energy were discussed in [9]. Efficiently benchmarking smart home challenges were explained [3]. And finally, a coordination platform for handling emergencies and restoration of power grid was introduced [5].

Organizing Committee

- Heba Mohamed, University of Bonn, Germany
- Nikola Tomašević, Institute Mihajlo Pupin, Serbia
- Marko Batić, Institute Mihajlo Pupin, Serbia

International Doctoral Committee

- Valentina Janev, Institute Mihajlo Pupin, Serbia (Chair)
- Sanja Vraneš, Institute Mihajlo Pupin, Serbia
- Lazar Berbakov, Institute Mihajlo Pupin, Serbia
- Emanuel Sallinger, University of Oxford, UK
- Anastasia Dimou, imec and Ghent University, Belgium
- Diego Collarana, Fraunhofer IAIS, Germany
- Maria-Esther Vidal, TIB, Leibniz University Hannover, Germany
- Jens Lehmann, University of Bonn, Germany
- Damien Graux, Inria, Université Côte d'Azur, CNRS, I3S, France
- Hajira Jabeen, CEPLAS, Technische Universität Dresden, Germany
- Andrej Čampa, ComSensus, Slovenia
- Marcus Keane, National University of Ireland, Galway, Ireland
- Dimitar Trajanov, Ss. Cyril and Methodius Univ, Skopje, North Macedonia
- Johannes Stöckl, Austrian Institute of Technology, Austra
- Federico Seri, National University of Ireland, Galway, Ireland
- Luis Miguel Blanes Restoy, National University of Ireland, Galway, Ireland
- Brankica Pažun, School of Engineering Management, Serbia
- Neven Vrček, Faculty of Organization & Informatics, Univ of Zagreb, Croatia
- Paulo Lissa, National University of Ireland, Galway, Ireland

Acknowledgements

We would like to thank all the authors, reviewers, committee members and the invited speaker for their contributions, support and commitment during this particularly challenging year.

This event was supported by the European Union Horizon 2020 project LAMBDA (Grant Agreement No. 809965). In addition, it received the sponsoring support of the Ministry of Science and Technological Development of the Republic of Serbia (No. 451-03-9/2021-14/200034) and the Science Fund of the Republic of Serbia (Artemis, No.6527051).

Articles presented at the Ph.D. workshop

1. Draschner, C.F., Moghaddam, F.B., Lehmann, J., Jabeen, H.: Semantic analytics in the palm of your browser. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
2. Hudomalj, M.: Traveling-wave event detection and localization on power cables. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
3. Jelić, M., Pujić, D., Batić, M.: Energy efficiency benchmarking for smart homes. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
4. Moghaddam, F.B., Draschner, C.F., Lehmann, J., Jabeen, H.: Semantic Web analysis with flavor of micro-services. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
5. Popadić, D., Batić, M.: Coordination platform for handling emergencies and restoration of power grid. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
6. Popović, M., Tomašević, N.: A blockchain-based platform for keeping logs of citizens' consents. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
7. Silva, J., Fraga, L., Teixeira, S., Teixeira, J.: Numerical tools developed to predict the combustion behavior inside a 20 kW pellet boiler. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
8. Sodin, D.: PMU-based fault localization in distribution networks. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
9. Stanković, K., Jelić, M., Batić, M.: The cloud-based control platform for multi-source renewable energy system. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)

References

10. Commission, E.: Building a European data economy (2017), https://ec.europa.eu/digital-single-market/en/policies/building-european-data-economy
11. Graux, D.: Hints to save time when dealing with Big Data. In: Proceedings of the 1st Ph.D. Workshop on Big Data Analytics from the LAMBDA Network (2021)
12. Graux, D., Janev, V., Jabeen, H., Sallinger, E.: A Big Data learning platform for the West Balkans and beyond. In: Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 2. pp. 617–618 (2021)
13. Graux, D., Janev, V., Jabeen, H., Sallinger, E.: Deploying a strategy to unlock Big Data research and teaching activities in the West Balkan region. In: Proceedings of the 26th ACM Conference on Innovation and Technology in Computer Science Education V. 1. pp. 491–497 (2021)
14. Janev, V., Graux, D., Jabeen, H., Sallinger, E.: Knowledge graphs and Big Data processing. Springer Nature (2020), https://link.springer.com/book/10.1007/978-3-030-53199-7
15. Khalili, A., Auer, S., Tarasowa, D., Ermilov, I.: SlideWiki: elicitation and sharing of corporate knowledge using presentations. In: Int. Conference on Knowledge Engineering and Knowledge Management. pp. 302–316. Springer (2012)
16. Zhou, A.C., He, B.: Big Data and exascale computing. In: S. Sakr, A. Y. Zomaya (eds) Encyclopedia of Big Data Technologies, Springer, Cham. (2019)