European Data Science Academy

Inna Novalija
Jožef Stefan Institute
Jamova cesta 39, Ljubljana, Slovenia

Erik Novak
Jožef Stefan Institute
Jamova cesta 39, Ljubljana, Slovenia

Marko Grobelnik
Jožef Stefan Institute
Jamova cesta 39, Ljubljana, Slovenia

inna.koval@ijs.si

erik.novak@ijs.si

marko.grobelnik@ijs.si

ABSTRACT

In this paper we present the results of the European Data Science Academy (EDSA) project that developed curriculum and training materials in the area of data science. EDSA dashboard became a tool for individuals, researchers and policy makers interested in data science demand and supply topics. EDSA partners produced a number of learning analytics instruments that contribute to the improvements of the learning process.

CCS CONCEPTS

• **Applied computing → Education**; • **Information systems** →*Information retrieval*;

KEYWORDS

EDSA, data science, demand analysis, dashboard, learning analytics.

1. INTRODUCTION

The 'Age of Data' is evolving, with data generated at a phenomenal rate that introduces numerous challenges regarding the collection, storage and analysis of this data.

The European Data Science Academy (EDSA) [1] was an H2020 EU project that ran between February 2015 and January 2018. The objective of the EDSA project was to deliver the learning tools that are crucially needed to close the skill gap in Data Science in the EU.

The EDSA project has developed a virtuous learning production cycle for Data Science, and has:

- Analyzed the sector specific skillsets for data analysts across Europe with results reflected at EDSA demand and supply dashboard;
- Developed modular and adaptable curricula to meet these data science needs; and
- Delivered training supported by multiplatform resources, introducing Learning pathway mechanism that enables effective online training.

2. Project Partners

The EDSA consortium with the Open University as a project coordinator was formed by 9 partners from 6 European countries (UK, Germany, Sweden, Slovenia, The Netherlands and France). Partners represented universities (The Open University, KTH Royal Institute of Technology, University of Southampton, TU/e Technische Universiteit Eindhoven University of Technology), research institutions (Fraunhofer IAIS, Jožef Stefan Institute) and SMEs (Open Data Institute, IDEXLAB SAS, Persontyle Limited) covering different forms of organizations.

3. Project Results

The prominent EDSA project result included the development of EDSA demand and supply dashboard that aggregates job vacancies in the area of data science, provides the insight into the skills required by data scientists and links the demand analysis results to supply materials.

Figure 1: List of jobs from EDSA dashboard.

EDSA partners developed a data science curriculum that helps modern data scientists to be on track with the necessary knowledge in the relevant fields.

EDSA project provided a number of learning analytics tools that allow for establishing more efficient learning mechanisms and processes.

3.1 Demand Analysis

EDSA has been monitoring trends across the EU to assess the demands for particular data science skills and expertise. Project partners have leveraged a vast network of European data providers, consumers and intermediaries to "track the pulse" of the European data landscape. This has allowed project partners to align the criteria with the latest demands of the community. EDSA partners have built an interactive dashboard [2] to present the current state of the European Data Science landscape, with the data feeding into the curricula development.

EDSA dashboard has been built upon two main components:

- Demand component that includes a job list in the area of data science with extracted skills, aggregated statistics for skills extracted from jobs, map with job clusters and

- Supply component that includes suggested courses in the area of data science, suggested videolectures [3] and the learning pathway that enables users to follow the dedicated data science curriculum.

Figure 1 presents a list of jobs in the area of data science from the EDSA dashboard, while Figure 2 illustrates the map and timeline for jobs.
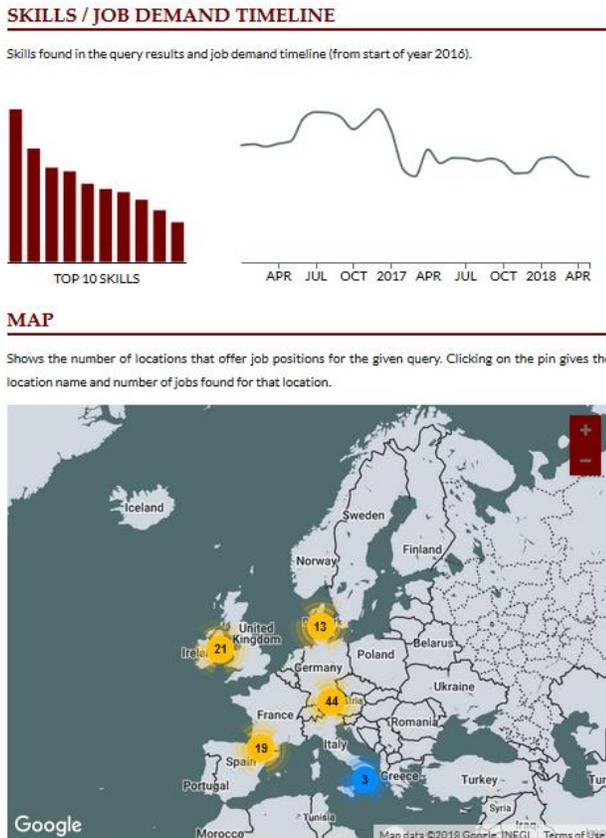
SKILLS / JOB DEMAND TIMELINE

Skills found in the query results and job demand timeline (from start of year 2016).

MAP

Shows the number of locations that offer job positions for the given query. Clicking on the pin gives the location name and number of jobs found for that location.

Figure 2: Jobs/skills map and timeline.

EDSA dashboard enables a search functionality based on job vacancies data semantically annotated with Wikifier [4] – a service allowing the identification and linking textual components (including skills) to the corresponding Wikipedia pages.

The learning pathway dashboard element (illustrated at Figure 3) provides to users the automatically built data science pathway according to user request.

3.2 Curricula Development

EDSA has developed a core data science curriculum based on topics extracted from the demand analysis. Project produced training materials to cover these topics, utilizing existing resources available in the public domain and the internal expertise of the EDSA consortium. All learning resources produced by the project adhere to the EDSA values and have been made available via the EDSA courses portal. The curricula design has arranged these resources into pathways that can be customized to suit the different needs of individuals, and the requirements of self-study and tutor-led training formats.

LEARNING PATHWAYS

The learning pathway in showing the courses that will give you the skills related to the query.
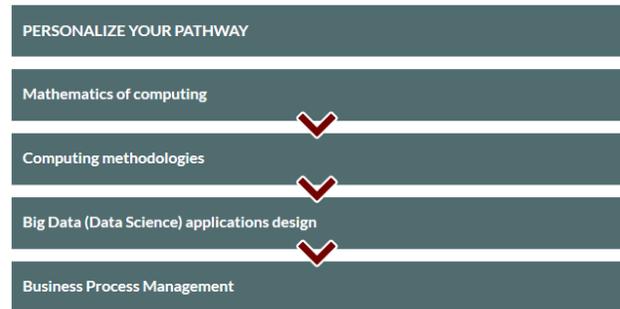

Figure 3: EDSA learning pathway.

3.3 Training Delivery and Learning Analytics

Key parts of EDSA curricula have been delivered through MOOCs, videolectures and face-to-face training. EDSA primarily used VideoLectures.NET portal from Jožef Stefan Institute and FutureLearn [5] – the largest European MOOC platform, founded by The Open University – to maximize outreach and uptake of project materials. Engagement with learners and Big Data stakeholders has been key to the training, and therefore monitoring and analysis tools have been utilized to assess learner progress – rather than by simply listening to academics or technology evangelists.

4. CONCLUSION

The European Data Science Academy project became a node that on one side explores the required skills in the area of modern data science and on the other side provides a possibility to obtain the necessary qualifications via training, tools and methods for learning analytics. Project partners agreed to maintain the developed results for at least one year after the end of the project.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] EDSA project, http://edsa-project.eu (accessed in May, 2018).

[2] EDSA dashboard, http://edsa-project.eu/resources/dashboard (accessed in May, 2018).

[3] VideoLectures.NET, http://videolectures.net (accessed in May,2018).

[4] JSI Wikifier, http://wikifier.org (accessed in May,2018).

[5] FutureLearn, https://www.futurelearn.com (accessed in May, 2018).