Project acronym:	**EDSA**
Project full name:	**European Data Science Academy**
Grant agreement no:	**643937**

D5.3 Preliminary individual, joint and external exploitation plans

Deliverable Editor:	**Jean-Louis Liévin (ideXlab)**
Other contributors:	**Simon Scerri, Angi Voss (IAS)**
	Ali Syed (Persontyle)
	Erik Novak, Inna Novalija (JSI)
	Mihhail Matskin (KTH)
	Chris Phethean (SOTON)
	Alexander Mikroyannidis (OU)
	Remi Brochenin (TU/e)
Deliverable Reviewers:	**Steffica Warwick, Leonard Mack (ODI),**
Deliverable due date:	**31/07/2016**
Submission date:	**31/07/2016**
Distribution level:	**Public**
Version:	**1.0**

This document is part of a research project funded
by the Horizon 2020 Framework Programme of the European Union

Change Log

Version	Date	Amended by	Changes
0.1	27/04/2016	Jean-Louis Liévin	Initial draft
0.7	08/07/2016	All partners	Inputs before consolidation
0.8	11/07/2016	Jean-Louis Liévin	Consolidation
0.9	25/07/2016	All partners	Revised, following reviewers' inputs and partners' amendments
0.91	25/07/2016	Jean-Louis Liévin	Revised, following partners additional input
1.0	29/07/2016	Jean-Louis Liévin	Final version

Table of Contents

Executive Summary

As the main collaborative route for the exploitation of the EDSA project is through the establishment of the European Data Science Academy (EDSA) as a body to lead data science training across Europe, this document complements the exploitation aspects project with individual exploitation plans.

The project is composed of three types of partners:

- Academic partners (the Open University, the University of Southampton, KTH, TU/e)
- Research organisations (Fraunhofer IAIS, the Josef Stefan Institute)
- SME partners (the Open Data Institute, Persontyle, ideXlab)

As was originally envisaged in the project proposal, the complementarities between the partners are reflected in the individual exploitation strategies envisaged at this stage.

The **academic partners** will seek to leverage the EDSA brand to increase the visibility of the developed curricula and educational resources: interactive eBooks and courses, MOOCs, etc.

The developed material will target various audiences: bachelor or MSc students and business professionals, and are intended to be delivered through various means, such as face-to-face, online, during summer schools, etc.

The **research organisations** envisage to further exploit their research outputs related to course analytics (e.g. through data mining, process mining) or course delivery methods (e.g. through innovative video lectures) as well as push and enhance their own training programmes (for post-graduates, business professionals, etc.)

The **SMEs** see their individual exploitation plans as exploring and developing new business opportunities, such as:

- Developing training programmes for new niches and currently unaddressed training needs.
- Focusing on content in high demand (e.g. machine learning).
- Reaching new customers through the EDSA Academy.
- Developing the coupling between training programmes and open innovation platforms and services.

All these individual exploitation plans will continue to be nurtured during the 2nd half of the project in conjunction with the development of the EDSA Academy exploitation.

Introduction

As described in the DoW work package 5, "WP5 develops an exploitation strategy for the outputs of the project based on the establishment of the European Data Science Academy (EDSA). WP5 will also coordinate individual partner and project wide strategic exploitation plans."

The present document highlights the individual partner exploitation plans.

The EDSA project exploitation strategy is aligned with the agendas of the individual beneficiaries. In order to demonstrate the value of the outputs of the project, the project partners have drafted exploitation plans that integrate the outputs of the project within their current business development processes. We list them below starting with our SMEs, then our research organisations and finally the universities.

For each partner, we start with a brief presentation of the partner activity, followed by a list of opportunity descriptions.

Each opportunity description contains the following paragraphs:

- Description: a high-level description (1-2 paragraphs description);
- Impact and relevance: Gap addressed by the opportunity and relevance to the EDSA project;
- Commercial aspects: IP protection/licensing if relevant and high-level description of commercial terms;
- Actions: Is the opportunity already seized and next steps beyond the project to further develop it;
- Timing: high-level tentative milestones.

The main collaborative route for exploitation is through the establishment of the European Data Science Academy (EDSA) as a body to lead data science training across Europe. EDSA will be founded by the project partners as core members and external organisations will be invited to become members of this entity. The EDSA exploitation route is documented in a separate document (D5.1: The EDSA Charter).

Individual exploitation plans

This section covers the individual exploitation plans of the partners. It contains a short description of the partner and its interest, as well as the opportunities the partner sees for exploiting results developed within EDSA.

Open Data Institute

Partner description

The provision of training is a core activity for the Open Data Institute (ODI). The ODI's Learning Programme has established a globally recognised programme of training courses, online content, events and lectures focusing on open data, data science and innovation.

The ODI will exploit the results of EDSA to inform the development of its data skills framework, curriculum and training content. EDSA is also providing opportunities to engage with the data science community across Europe.

To further spread the uptake of new developments and support community engagement, the ODI will draw upon its established network of nodes. ODI Nodes are franchised organisations, authorised to use ODI products. Eleven ODI Nodes focus on learning activities and deliver ODI training across Europe and globally. Disseminating EDSA derived insight and content through the Node network will help extend the impact of the project.

The ODI is delivering data science training as part of the EDSA project. Courses include Open Data in Practice, Finding and Preparing Data, Visualising Data, and Open Data Science. EDSA will directly shape the development of new curricula both for EDSA and in other ODI training forums. As well as public training, the ODI offers tailored in-house training, which is becoming increasingly popular with multinational commercial organisations. The educational resources and delivery channels of the EDSA project will be essential to ODI for fulfilling the requirements of the training activities and it is anticipated that they will be used extensively. A list of public ODI courses can be found on the courses page of the ODI website[1].

Opportunity 1: Improving the content and delivery of ODI training

Description

Building on the EDSA demand analysis, we have identified new niches and currently unaddressed training needs to improve the content and delivery methods of ODI training. Insights gathered through the analysis are informing the revision and development of new courses - particularly online training, workshops and webinars. Following insights derived from the demand analysis, we will focus on creating blended training approaches. Blended training will combine online, classroom and virtual options to ensure both an effective learning experience and scalability.

We aim to build on insights from the EDSA project work by:

[1] theodi.org/courses

- **Integrating data science specific soft skills training.** The ODI is exploring options to extend its curriculum through soft skills training as advised by the EDSA's demand analysis.[2] Our goal is to answer the question in what ways are classic soft-skills, like presenting, different for data scientists. - and to find the most effective ways to develop these skills in data scientists.
- **Using the ODI Skills framework to simplify learning journeys.** The ODI is currently piloting a skills framework to structure flexible journeys to data literacy. The goal is to provide learners with an easy to use tool to help them identify the skills they need and build a learning journey that suits them.
- **Providing data literacy training.** Curriculum modules and feedback gathered through the EDSA will help to catalyse and further shape the ODI's offer for data literacy training. Data literacy has many dimensions, including ensuring that those commissioning and benefitting from data science activities understand enough to make the right decisions and investments. The output can be expected to include content and courses for use outside of data science teams.
- **Developing courses and content laterally related to EDSA:** This will include the development of new modules on "Finding stories" and "Licensing and data markets".
- **Creating and curating flexible content types:** Guided by the demand analysis, we will identify priority content areas and create content to meet learner needs. The content will be designed to be flexible, meaning it can be incorporated into different kinds of learning journeys. For example, a case-study that can be used as a standalone piece of content, but that also has some associated lecture notes so that it can be used by a trainer in an ODI node as part of a workshop.
- Improving learner experience. We will include insights from EDSA's learning analytics statistics in order to improve the delivery and learning experience for ODI training courses.

Impact and relevance

The ODI's expected impact from this activity is to increase the ROI on data science projects by improving the way that data is understood across all levels of organisations. The ODI has delivered training to more than 5,000 people from the public and private sectors, thereby fuelling data-driven innovation and making Europe more competitive. In this context, the expansion of the ODI's course and content offer is important to continue serving industry demands as effectively as possible.

Commercial aspects

The ODI generates revenue from training courses in multiple formats, including public and bespoke face-to-face training and via development of customised online material for clients. The ODI develops free and openly available online content, which leads to participants signing up for paid-for training. New courses developed as a result of the EDSA project will follow the same revenue model.

Actions and timings

The following high-level actions will need to be taken:

- Design and test of ODI soft skills training module (Q4 2016).
- Delivery of soft skills training pilot (Q1-2 2017).
- Development and test of ODI skills framework (in progress).
- Revision and adaptation of EDSA data literacy training to align with ODI skills framework (Q4

[2] Please refer to D1.4, currently under review and unpublished. After review it will be available via the EDSA project website: http://edsa-project.eu/downloads/deliverables/

2016).

Opportunity 2: EDSA Charter - new models for the delivery of data science training

Description

Through the EDSA Charter, the project consortium seeks to establish a long-term training offer through a network of partners. Therefore, the consortium will explore and pioneer new business and franchise models for the co-delivery of training through a federated, open network. The ODI aims to capitalise on this work, especially to further develop the ODI's global learning network. The ODI will use the EDSA Charter model to broaden the ODI's learning network and to test the model's practical feasibility.

Impact and relevance

Co-delivery of training through a network of franchise partners will scale and expand skills training across regions. For the ODI's open data and data literacy training, this approach is particularly relevant in order to deliver affordable and effective training in other countries than the UK.

We anticipate that the training approaches pioneered by EDSA will deliver new insights and practical learnings for how cross-country learning networks should be structured and which delivery modes are effective.

Commercial aspects

EDSA consortium members and external partners will need to establish the business models and commercial structures that will sustain EDSA beyond the term of the project. This will need to include agreement on what EDSA will offer as a minimum and strategies for its development. Partners will also need to agree details of associated legal structures, financial processes and governance models. The details of these arrangements will need to be documented, agreed upon and signed by all participants.

Actions and timings

The following high-level actions will need to be taken:

- Establish EDSA structures (business, legal, financial and governance) (Q2-3 2016).
- Document EDSA structures and partner agreements in EDSA Charter (Q2-3 2016).
- Engage potential partners to help develop agreements (Q3 2016).
- Sign-off final agreement on EDSA Charter by all project partners (Q4 2016).
- Establishing and kick-off of Academy (Q1 2017).

Opportunity 3: Increase ODI's exposure and network within the European data science community

Description

Through the EDSA project, the ODI has accessed a wide pool of data scientists and their managers from across Europe. This has substantially increased the ODI's exposure to new communities and commercial sectors, offering an opportunity to expand the ODI's European network. Through this exposure and by developing a continued relationship with data scientists and their managers from a variety of small, medium-sized and large companies, we anticipate new business development opportunities, e.g. by delivering customised trainings.

Impact and relevance

While the ODI has an extensive global network, there are many regions and industry sectors within Europe where our network would benefit from additional and stronger relationships. This is particularly true for the data science sector. The ODI has developed and is planning further development of data science courses as well as training addressing the soft skills gap identified as part of the EDSA demand analysis. An increased network among data scientists will provide the initial audience for these courses. Additionally, we plan to explore new options for delivering data science and data literacy training customised to various industry sectors. This ongoing work is also in line with EDSA's main project ambition of closing the data science skills gap across Europe.

Commercial aspects

All individuals that took part in the demand analysis were asked if they wished to be contacted about the project's progress and separately, if they wanted to be sent the results of the demand analysis. This provides one basis for further engagement with the network.

There is direct commercial benefit from building and maintaining the ODI's network, for example by generating revenue through course attendance if these are successfully promoted through this network.

Actions and timings

The following high-level actions will need to be taken:

- Contacting demand analysis study participants (Q3 2016).
- Update study participants on results, impact (in terms of curriculum adaptation), and general project progress (ongoing after Q3 2016).
- Explore options with participating organisations on further training needs and potential delivery (ongoing).

ideXlab

Partner description

ideXlab is an ambitious start-up focused on open innovation. More specifically, ideXlab is like a meet.com between innovative companies and hard-to-find expertise. As even the largest companies cannot keep up with the pace of the technology evolution in every domain, ideXlab has developed a unique solution to pinpoint for them the most specialist experts around the world, and make the matchmaking easier. Its technology platform analyses more than 200 million technical documents and can reach more than 10 million academic experts around the globe. These numbers will continuously increase over the next few years. Its core market – innovative enterprises – comprises more than 1 million companies worldwide.

The ambition of ideXlab with EDSA is to couple online learning resources with open innovation, therefore gaining the capability to offer and push educational resources in addition to providing access to expertise.

Opportunity 1: Coupling between open innovation and online learning resources

Description

ideXlab is about technologies and strategies for open innovation and on instruments to mediate between

industry problems and expertise available in the market. Data science covers a set of skills that are becoming increasingly relevant within industry. It is therefore an ideal test bed to investigate new business opportunities, in particular expanding the existing ideXlab product portfolio through new services relying on online learning resources. Online learning resources can become a natural extension of open innovation by offering to innovating companies, in addition to intermediation services with experts, some educational resources to upskill their employees.

Impact and relevance

Companies, and SMEs in particular, use open innovation mechanisms and platforms to gain access to external expertise to help them solve internal problems. Bringing in the right knowledge and skills to address recurrent problems is essential to ensuring long-term competitiveness. In this sense online learning resources can become a natural complement to open innovation tools: while open innovation provides easy access to external expertise, online learning resources provide an efficient mean to increase internal expertise. In the project, ideXlab was particularly interested in the relationship between its current open innovation core business (as an intermediary between companies and external experts) and training material. ideXlab will extend the current algorithms which are used to identify and rank experts according to a question submitted by an enterprise in order to also cover and identify online learning resources such as MOOCs. These resources could be used to train in-house staff.

Commercial aspects

The EDSA (Academy) Terms and Conditions will apply, as well as the usual ideXlab Terms and Conditions.

Actions

Further investigation is required to assess the coupling aspects between open innovation and online training resources such as:

- Service mock-up based on data science
- Socialisation with early customers
- Investigation of extension opportunities to other domains

Timing

During the last year of the EDSA project, ideXlab will envisage and test how the coupling may apply, and how it might be extended from data science to other disciplines. This will be developed through an initial Beta service implementation that will be socialised with a set of early customer for feedback, improvement and need assessment, beyond data science.

Opportunity 2: Reaching new customers

Description

Through its contributions to the EDSA project, ideXlab will engage with new prospects having specific needs in data science training (e.g. CIOs of French R&D groups). Such companies will have requirements and strive to collaborate with research entities and other companies to solve specific big data problems. The model may also be applicable to other domains in high demand (e.g., biomimetics, additive manufacturing, smart materials, Internet of Things, etc.) where course material would lead to developing new business opportunities.

Impact and relevance

This activity will contribute to the promotion and dissemination of the EDSA project outputs. In particular, it will help to expand the user base of both EDSA's and ideXlab's products, thereby growing a sustainable client base.

Commercial aspects
The EDSA Terms and Conditions will apply, as well as the usual ideXlab Terms and Conditions.

Actions
ideXlab will re-use and promote whatever dissemination material will be made available by the project or the Academy:

- Webinar and blog post on links between open innovation and online educational material
- Webinar and blog post on data science state-of-the-art as found by ideXlab platform and EDSA educational resources
- Promotion of the above through face to face and conferences opportunities

Timing
The activity will start immediately when EDSA Academy material is made available.

Persontyle

Partner description
Persontyle has already developed a comprehensive set of data science and engineering learning opportunities for people to develop better insights and deliver meaningful value using data. Persontyle offers instructor led courses, bootcamps, and open learning programs that provide a hands on learning environment, training labs designed to cover real world use cases, and applied knowledge of the data science practices and tools.

Persontyle, through its education initiative called "The School of Data Science", is focused on spreading and democratizing access to big data, data science, data engineering and machine learning skills. Please note that "The School of Data Science" is not an accredited institution (i.e., we do not award degrees).

Persontyle will exploit the results of the project by refining and balancing its current data science and engineering, machine learning and advanced analytics curricula and plans to offer on demand eLearning programs and specialized industry focused training. Persontyle, in collaboration with EDSA, will design and deliver open learning events and workshops across Europe for business and technology professionals to develop new data skills. Additionally, Persontyle will leverage the output of the demand analysis work to design new data science training programs.

Opportunity 1: Developing a machine learning study guide
Develop and publish a study guide for anyone who wants to learn practical machining learning using R.

Description
Develop a comprehensive guide with 19 modules focused on a practical introduction to machine learning and predictive models, which is intended to serve as a fundamental resource for data scientists, analysts and programmers. The guide will help develop an applied understanding of the principles of machine learning and enable them to develop practical solutions using predictive models.

Impact and relevance

There is an unfortunate air of mystique that surrounds the advanced machine learning techniques. Too many data scientists and data science stakeholders fear that actual deep understanding of the algorithms is beyond them. This is a serious problem, for without this understanding it is impossible to consistently do good work within data science. Nor is it possible to explain why what succeeds did succeeds – or, sometimes more pressingly, why what will succeed will succeed.

This guide aims to show readers that they can understand these techniques. For it is a wonderful truth that the advances of machine learning that are changing the world are no more than the layering of a number of simple techniques. Once you see what these simple techniques are and how they fit together, their patterns are clear and the mirage of complexity dissolves.

Commercial aspects

Actions

Publish the final version of the guide with including following four new modules added:

- Gaussian Processes
- Tree based methods
- Support Vector Machines
- Neural Network and Deep Learning

Timing

First version of study guide already published. Version 2.0 to be released by December 2016.

Opportunity 2: Develop new Machine Learning MOOC Program (MOOCs) in collaboration with FutureLearn and OU

This flexible online machine learning program will enable learners to master the subject of machine learning in depth — ideal for developing career or preparing for further study.

Description

Though it has been an area of active research for over 50 years, machine learning is currently undergoing a renaissance driven by Moore's law and the rise of big data. Large private and public investment in the area has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome.

In partnership with EDSA and FutureLearn, we will develop a practical programme that uses hands-on examples to step through real-world application of machine learning. This will enable participants to understand the basic concepts, become confident in applying the tools and techniques, and provide a firm foundation from which to explore more advanced methods.

The aim of this programme is to provide both a deep understanding of the techniques and practices of machine learning and to expose a wide set of resources capable of being wielded by the data scientist and analyst in their work. Participants will encounter explanations of the theory behind the algorithms and models they are exposed to, giving them an understanding of the strengths and weaknesses of each which they should be able to use to reason about suitable approaches to real life problem – and to

communicate such reasoning to other stakeholders in such problems.

Programme outline:

- Course 1 - Machine Learning Essentials
- Course 2 - Basic Machine Learning Methods
- Course 3 - Machine Learning using Open Source and Cloud Platforms
- Course 4 - Advanced Machine Learning Methods

Impact and relevance

This programme is aimed at a number of audiences. These include professional data scientists, data analysts, developers, and those aiming to become such, advanced undergraduate and postgraduate students, and researchers from areas outside data science looking for a guide to the utilization of these techniques in their work.

Machine learning is a core element of data science and by attending this program participants will be able to understand the art and science of discovering patterns and making intelligent predictions from big data. They will also be able to define machine learning, articulate why it matters, and discuss its relationship to analytics, data science, and big data. Participants will practically learn the most commonly used machine learning methods, covering both supervised and unsupervised learning.

Commercial aspects

Actions

- Finalise the contract with OU.
- Design and development of content for course 1 and course 2.

Timing

We aim to launch the first of the course program by January / February 2017.

Opportunity 3: Professional certification program

Design and develop an outline of the professional certification program based on the inputs of the demand analysis and inputs from other EU projects like Edison.

Description

The certification program will help professionals and job seekers to demonstrate their big data and data sciences expertise and gain recognition for one of the most sought after skills in technology today.

Impact and relevance

There is strong demand for accredited and certified data scientists across Europe. Once designed and launched, candidates who meet the preparation requirements and pass a certification exam will receive many benefits:

- Industry recognition for widely sought after data science skills;
- official designation and logo that can be used on business cards and online profiles;
- digitally verifiable credential for employers and clients.

Commercial aspects
License terms for this work need to finalised as part of the certification program design work.

Actions
- Work with EDSA and EDISON project members and publish a broad outline of the certification program by March 2017.
- Develop a high-level structure of the professional certification program by June 2018.

Timing
Draft of the certification program to be published by June 2018.

IAIS

Partner description
Through its Academy, Fraunhofer offers outstanding courses of study, certificate courses and seminars to specialists and managers based on the research activities of the Fraunhofer institutes. Fraunhofer IAIS will exploit the project results for the "Data Scientist Training" of the Fraunhofer Gesellschaft, which targets professionals with face-to-face trainings in small groups.

Since 2013, IAIS has given face-to-face courses for professional data scientists in different sectors and on different topics. These courses have been extended to a training programme in the Fraunhofer Alliance "Big Data" in 2015 with complementary courses from other Fraunhofer Institutes, partially supported by the professional school of EIT Digital. This data scientist training programme became part of the Fraunhofer Academy's professional training offerings. In 2016 a certification programme of three stages was added. At the first and most basic stage, a five-day course called "Data scientist basic level" is offered, which is a joint effort of four institutes. At the second stage, certificates will be developed for advanced courses on specific topics or for specific sectors.

IAIS also hosts the multilingual open education platform SlideWiki.org, which is used for a number of eLearning activities.

Opportunity 1: Enhancing our data scientist courses
Description
With our contributions to EDSA we want to enhance our face-to-face courses. This can be done in several ways: as preliminary information for self-assessment before someone registers to a face-to-face course, to prepare for the face-to-face course, to prepare for an exam associated with the course or, as a take-away with appetizers for further courses. As a side-effect of our engagement, we hope to raise the reputation of our courses.

Impact and relevance
Face-to-face courses are expensive. Therefore, persons interested in our courses require information to make sure that the selected course is relevant for them and justifies the expense. This information can be provided by online learning resources that we submitted to EDSA.

Occasionally, participants of our courses ask for preparatory material and we want to make sure that they can follow the course. Recommending EDSA material can serve both purposes.

For courses which can be combined with a certificate, participants require additional summaries and exercises. This could also be provided by the online learning resources that we submitted to EDSA.

Commercial aspects

We understand that we are giving away our learning material under the open source license used by EDSA. Therefore, we do not publish the full content of our courses to EDSA.

Actions

We have already used EDSA learning resources as preparatory material for one of our courses. We stopped it as about half of the participants could not access the EDSA courses, which require an account at Google, Linked-In or Facebook. We would need a way to self-register with the EDSA portal.

To use EDSA material for advertisement and self-assessment before registration will require a management decision, which might depend on the popularity gained by EDSA.

Any summaries and exercises which will be developed for the courses in our certification program would have to be adapted and made available as online material. A concrete option was discussed with EIT Digital at the beginning of this year. Depending on their success with blended courses in their professional school and depending on the success of our certificate course EIT Digital would invest in developing corresponding online material.

Timing

- Q4 2016: Once EDSA allows participants to register, continue to provide EDSA learning resources for participants to prepare to face-to-face courses.
- Q4 2016: Management decision for using EDSA material for advertising and self-assessment.
- Q1 2017: Development of online material for our course “data scientist basic level” with EIT Digital for their professional school.

Opportunity 2: Providing a test bed for innovative data analysis methods

Description

The EDSA use-case - analysis of job posting trends for demand monitoring - provides for a very valid test-bed for the data analysis methods being developed by Fraunhofer IAIS.

Impact and relevance

Data mining algorithms and methods for clustering and trend analysis can be applied to this real-world application, by taking the extracted job posts as input.

These methods will complement the (manual) exploratory data analysis functionalities provided by the EDSA dashboard and provide further insights based on the results of fully-automated analysis. The latter are being developed in separate efforts being undertaken by Fraunhofer researchers. The methods, which can operate on any time series, will in this case be integrated in the dashboard to produce insights for the EDSA use-case, i.e., gauge trends and project future directions in skill demands based on past data.

Commercial aspects

The re-use of the data analysis methods tailored for the EDSA use-case will be bound by the EDSA agreement. Commercial use by third parties is also possible within that context but appropriate remuneration to Fraunhofer IAIS is expected.

Actions

The methods will be integrated in the EDSA dashboard to showcase their value for this use-case, as a proof-of-concept.

Timing

- *Up to M18: Ensuring that the results of the EDSA Acquisition Task conform with the input required for the Data Analysis methods.*
- *M18-M24: Tailoring the existing method being produced by IAIS for the EDSA use-case*
- *M24 or later: Integrating the methods in the EDSA dashboard*

JSI

Partner description

JSI will exploit project results along three main lines:

1) Through the VideoLectures.NET portal.

2) By extending the PhD and MSc programs at "Jožef Stefan International Postgraduate School".

3) Through the "Knowledge For All" foundation.

VideoLectures.NET is an award-winning free and open access educational video lectures repository. The lectures are given by distinguished scholars and scientists at the most important and prominent events like conferences, summer schools, workshops and science promotional events from many fields of science. The portal currently offers over 20k of educational videos distributed through 579 categories.

Jožef Stefan International Postgraduate School (IPS) was established in 2004 as an independent higher education institution. Its study programmes were approved by the Slovenian National Council for Higher Education. The initiative for the establishment of IPS came from the Jožef Stefan Institute (JSI). It was strongly supported by industry (Gorenje, Kolektor, Salonit) and an international network of cooperating universities and research institutes from the European Union, the USA, Japan, and a number of other countries. Jožef Stefan Institute provides the central research-educational basis. Within the IPS invited research institutes, industrial and other enterprises contribute their knowledge and innovation capacities for solving developmental problems.

Knowledge for all Foundation[3] is an initiative that brings together an alliance of established

[3] http://videolectures.net/k4a/

international institutions into a forum for discussion and dissemination of advances in innovations in technology enabled education at the national level in established and emerging nations. The foundation enables the provision of online academic videos, papers, technical reports, software and other materials, and actively promotes their free accessibility. Through its activities, the foundation will promote the spread of knowledge and rigorous academic debate to anyone with access to the internet anywhere in the world.

Opportunity 1: Enhancing the operability of the VideoLectures.Net portal

Description

Through the VideoLectures.NET portal, the goal is to structure thousands of video lectures within the proposed curricula as well as integrating learning analytics mechanisms to follow the progress of the student population.

Using applications developed within EDSA demand analysis task, we aim to connect demand for data science in Europe with supply of training materials available at the VideoLectures.Net portal.

Impact and relevance

By integrating additional functionality and dissemination we aim to increase the audience of the VideoLectures.NET portal from approximately 15,000 unique users per day towards a few tens of thousands.

Commercial aspects

Unless otherwise noted, Data Science videolectures presented on VideoLectures.NET portal use Creative Commons Attribution Non-commercial No Derivative Works licenses.

JSI applications developed within EDSA project are based on BSD licenced tools.

Actions

Since the start of EDSA project, we have over 30.000 views of Data Science videos published in 2015-2016.

We are developing the connection between the applications from demand analysis task and VideoLectures.NET portal. In particular, we would like to provide recommendations for each video based on the job seekers view (http://jobs.videolectures.net/jobseekers) of EDSA demand analysis dashboard.

Timing

Already started: Integration of demand analysis applications with VideoLectures.NET portal (active work until M36, prototype available in M24).

Already started: Recommendations for data science videos from VideoLectures.NET portal at EDSA website. More details can be found on the following link: https://edsa-project.eu/video-lectures (active work until M36, first results already available at EDSA website).

Already started: Learning analytics tools for VideoLectures.NET portal (active work until M36, first results already available, described in EDSA deliverable D3.3).

Next to be done: Support of connections between demand for data science professionals and data science training materials (start date: M19, completion date: M36).

Next to be done: Further increase the audience of the VideoLectures.NET portal (start date: M19, completion date: M36).

Opportunity 2: Extending the PhD and MSc programs at "Jožef Stefan International Postgraduate School"

Description

Extending the PhD and MSc programs at "Jožef Stefan International Postgraduate School" by introducing additional, specialized data science courses supported by high quality online content and services.

Impact and relevance

Data Science courses developed through the EDSA project fill the content gap for the existing courses at the "Jožef Stefan International Postgraduate School". EDSA learning analytics mechanisms can be used for triggering the better performance of students.

Commercial aspects

Unless otherwise noted, JSI training material developed within EDSA project have Creative Commons Attribution Non-commercial No Derivative Works license.

JSI applications developed within EDSA project are based on BSD licenced tools.

Actions

JSI developed training materials for EDSA online course "Foundations of Big Data". JSI is planning the following exploitation actions for the opportunity 2:

- extension of the "Jožef Stefan International Postgraduate School" courses with available EDSA training materials;
- providing recommendations to students about available EDSA trainings - online, blended and face-to-face courses.

Timing

Already started: Developed materials for EDSA online course "Foundations of Big Data" (completion date: M12).

Already started: Developed curriculum for EDSA online course "Statistical/Mathematical Foundations" (completion date: M18).

Next to be done: Provide recommendations about available EDSA trainings to the students of the "Jožef Stefan International Postgraduate School" (start planned for: M19, completion date: M36).

Opportunity 3: Internationalization through multilingual services

Description

Exploitation through the "Knowledge For All" foundation with the primary goal of internationalization through multilingual services and dissemination in non-English speaking parts of the world.

Impact and relevance

The multilingual demand analysis data obtained through the EDSA demand analysis and connected to VideoLectures.NET portal allow to fill the existing internationalization and dissemination gap.

Commercial aspects

Unless otherwise noted, JSI training material developed within EDSA project and Data Science video lectures have Creative Commons Attribution Non-commercial license.

JSI applications developed within EDSA project are based on BSD licenced tools.

Actions

The multilingual nature of data within EDSA project allows for internationalization and exploitation of achieved results via "Knowledge For All" foundation. The VideoLectures.NET portal itself contains a set of multilingual options. JSI actively performs dissemination of all achievements in this area:

- direct dissemination through VideoLectures.NET portal;
- dissemination through the events covered by "Knowledge For All" foundation.

Timing

Already started: Work with multilingual data. In particular, JSI started the multilingual data analysis within demand analysis task and integration of the results into VideoLectures.NET portal (completion date: M36).

Next to be done: Active dissemination of achieved results (completion date: M36).

The Open University

Partner description

The Open University (OU) is a major provider of free and open online learning materials worldwide, reaching more than 65 million downloads on Apple's iTunes U.[4] Most downloads originated from the English-speaking countries, closely followed by other European countries. While some of the eBooks are free, many are offered at an affordable price, contributing to the sustainability of the undertaken educational activities. OpenLearn,[5] an OER repository from the OU, is used by over 4.5 million users each year and has had over 38 million learners since its launch in 2006. The site averages 400,000 unique visitors a month with over 10,000 hours of learning materials including 8,000 hours taken from undergraduate and postgraduate modules. The OU's Stadium[6] facility supports a vast range of educational webcasts within the OU and external clients. Since its launch in 2000, Stadium has hosted a total of 1,488 public webinars. On average, Stadium receives approximately 2,000 visits per month and 1,400 unique visitors. The OU has a strong relationship with the BBC and co-produces a number of highly popular science programs such as COAST,[7] Bang Goes the Theory[8] and Frozen Planet.[9] The BBC

[4] http://projects.kmi.open.ac.uk/itunesu/impact/

[5] http://www.open.edu/openlearn/

[6] http://stadium.open.ac.uk/webcast-ou/

[7] http://www.bbc.co.uk/programmes/b006mvlc

[8] http://www.bbc.co.uk/programmes/b00lwxj1

[9] http://www.bbc.co.uk/programmes/b00mfl7n

programme "Don't Panic: The truth about population" featuring Hans Rosling, which forms the basis for a forthcoming FutureLearn MOOC on "Data visualisation for Development", had over 1.2 million viewers.

Opportunity: Increased visibility of the EDSA learning materials

Description

The OU will integrate the developed curricula and educational resources into the courses taught at the OU. The project's educational resources will be reused, repurposed and further developed by existing and emerging OU courses. Following the established tradition of open learning, the resources will be offered via iTunes U and our open institutional learning management system (LMS). Interactive eBooks and courses will be published based on the project's learning resources, delivered through a wide range of platforms, including iOS and Android devices, as well as desktop computers. Suitable material will be added to our FutureLearn MOOC offerings and project results will also be taken up by periodic summer schools and various seminars organised by OU researchers across the world including the ESWC Summer School.

Impact and relevance

The high quality EDSA learning materials will be a valuable addition to the relevant courses offered by the OU. Additionally, EDSA will benefit from the increased visibility of its learning materials, as these will be made available via the worldwide renowned educational channels of the OU.

Commercial aspects

The EDSA learning materials will be distributed via the OU educational channels using Creative Commons licenses, thus allowing the reuse, repurposing and republishing of the materials. Commercial use of the materials may also be allowed depending on the used licences.

Actions

A first version of the EDSA eBook has already been made available via the EDSA Online Courses portal. This eBook will be updated throughout the duration of the project and will be published by the OU on iTunes U as an open interactive eBook or a collection of open interactive eBooks.

The EDSA learning materials have also been taught by OU tutors in last year's ESWC Summer School and will be taught again in the next versions of this annual event.

Timing

There will be quarterly updates of the EDSA eBook(s) throughout the duration of the project and publication of the eBook(s) via the OU channel on iTunes U during the last year of the project. The eBook(s) will remain in iTunes U as a sustainable result beyond the end of the project.

The EDSA learning materials will be taught in the annual ESWC Summer School throughout the duration of the project and beyond.

SOTON

Partner description

SOTON delivers a range of higher education courses at Bachelor, Master and PhD level, and has been

leading in developing MOOCs for the FutureLearn platform since its release. The WAIS group at Southampton has an excellent marketing output, with projects and initiatives of the group (e.g., in the areas of linked data, open data, web and data science) regularly featured in highly visible programs and media channels, including the BBC, Financial Times, Times, Wired, and Communications of the ACM.

SOTON will exploit project results for both educational and industrial purposes. The educational material for cross-border and cross-sector data analytics produced in this project will contribute to the training of graduate students in a new data science MSc programme, as well as through Southampton's cross disciplinary doctoral training centre that has been funded to train 90 students for four-year Web Science PhDs. SOTON will also disseminate the training materials for data management courses provided by the Administrative Data Research Centre for England (ADRC*e), which SOTON leads, for different levels of researchers. On the industrial side, the Research and Innovation Services in SOTON has a core mission to promote and support the transfer of research expertise and knowledge into industry through training and collaboration with industrial partners, maximizing the commercial and societal impact of research output. The IT Innovation centre, as part of the School of Electronics and Computer Science in SOTON, also has extensive experience to apply new technologies from the research community to problems in industry and commerce. Both departments above in SOTON will: 1) ensure that the dissemination and communication activities in WP4, which SOTON leads, will maximize its outreach in industry; and 2) disseminate and promote EDSA results to a wider range of commercial audiences.

Opportunity 1: Delivering data science graduate courses at the University of Southampton

Description

Expanding the graduate programmes on offer at the University of Southampton, we have begun providing an MSc in Data Science since September 2015. This includes two new face-to-face modules related to the modules we have released within EDSA: Foundations of Data Science and Data Visualisation. We are now also in the process of planning Data Science 'conversion' courses having been awarded a HEFCE grant, allowing graduates from disciplines such as Physics and Chemistry to take modules that can equip them with the technical skills required for studying data science.

Impact and relevance

Data science courses are varied and can focus on numerous different elements of the wider discipline. By basing our MSc programme on the EDSA curriculum, particularly drawing on our experience developing the curriculum for Foundations of Data Science, we ensure that a broad coverage of the data science pipeline is provided to students, who can then choose to specialise their knowledge on a number of optional modules covering topics such as data mining and machine learning. On the conversion courses, students can choose from a mix of modules within their original discipline and from within the department of Electronics and Computer Science, allowing them to combine their new technical knowledge with continued development of their own subject expertise and domain knowledge.

Commercial aspects

Learning materials are being developed as part of the EDSA project and are therefore disseminated openly - the course slides are therefore available as self-study learning resources through the EDSA course portal.

Actions

The Data Science MSc launched in September 2015, making use of the curricula developed for Foundations of Data Science and Data Visualisation in WP2. The development of the Data Science conversion courses is now currently in progress. The first addition, to be developed Q3/Q4 2016, will be a module through which to teach graduates of all backgrounds the technical and programming skills required to begin to undertake data science training. Upon completion we will offer a range of Data Science MSc courses to graduates from a range of disciplinary backgrounds, meaning that they will be able to join a course which will teach them the discipline-specific techniques about data analysis along with the general programming and computational knowledge that will allow them to take advantage of big data technologies and cloud computing.

Timing

Already started: Development of Foundations of Data Science (ongoing since Q2 2015) and Data Visualisation courses (ongoing since Q3 2015).

Next to be done: Development of programming and computational thinking module for EDSA curriculum (commencing Q3-4 2016), to be used in the conversion courses.

Opportunity 2: Professional data science training

Description

In conjunction with the graduate courses described above, we are developing four Continuing Professional Development (CPD) courses in collaboration with the Cambridge Education Group (CEG). These will be online-based professional paid-for courses that will target those workers wishing to upskill into data science areas.

Impact and relevance

Providing training that is accessible to the public regardless of their location or circumstances is essential in order to allow anyone to acquire the skills necessary to become a data scientist. By offering high-quality, paid-for training targeted at professionals, rather than graduates, we ensure that we cover a broad range of audiences.

Commercial aspects

As professional CPD courses, these will be charged for and revenue will be split between the University of Southampton and CEG.

Actions

The CPD courses are currently in development, with the first (data science fundamentals) due to launch in November 2016. The current development schedule for the remainder of the courses is as follows:

Data science fundamentals:	November 2016
Data science for marketing:	Q1 2017
Data science for finance:	Q2 2017
Data science for health care:	Q3 2017

Timing

Already started: Development of first CPD in fundamentals of Data Science (ongoing since Q2 2016, to be released Nov 2016).

Next to be done: Deployment of first CPD course (Nov 2016, repeating), and development of three other domain-specific variants (Starting Q1, Q2 and Q3 2017 respectively, and then ongoing).

Royal Institute of Technology (Kungliga Tekniska Högskolan – KTH)

Partner description

Education and training is an essential part of KTH activity as an academic institution. KTH is Sweden's biggest technical university, delivering education to more than 14,500 Bachelor and Master students, as well as 1,700 doctoral students. Additionally, KTH provides tailored corporate training, commissioned education and further education for professionals. KTH is a core partner of EIT Digital; it participates in education in both the Masters and Doctoral schools.

KTH will exploit results of the project via the coordination of several masters programs:" Software Engineering of Distributed Systems" at KTH, the "Data Science" program of EIT Digital and the "Data Intensive Systems" specialization of the EIT Digital Master program "Cloud Computing and Services". Additionally, KTH will work on the incorporation of project results into relevant Erasmus Mundus programs (both Master and Doctorate) and into a set of commissioned courses given to external industrial organizations.

Opportunity: Quality improvement and distribution of KTH learning material

Description

KTH considers the work on the EDSA project as a part of its total efforts in enhancing learning materials and training opportunities in the area of Data Science. Spreading online materials and information about the given courses via EDSA extends the awareness about KTH efforts in this area. Exploiting of the EDSA consortium's training approaches leads to a better standardization of the delivered material, learning analytics and delivery channels. The main opportunity for KTH with respect to exploitation includes contiguous improvement of the quality of produced learning materials, exploitation of the developed training delivery channels, usage of EDSA demand analysis tools for adjustment of the training strategy to the real market needs and joint development of the training through both existing (for example, together with TU/e in the EIT Digital Master programs) and new channels.

Impact and relevance

The gap addressed by the opportunity is a possibility to build learning materials and perform training not only from the general university education perspective but also by adjusting them to dynamic demands from markets and provide opportunities for professions to get access to high level education materials in the relevant subjects.

Commercial aspects

We assume that IP rights in the exploitation of KTH results in EDSA will be based on the rules and regulations following from the Swedish law and the EDSA Consortium Agreement.

Actions

Changes to the curricula of Master programs "Software Engineering of Distributed Systems" (KTH) and "Cloud Computing and Services" (EIT Digital Master program) were performed, taking into account EDSA demand analysis results. Further adjustments of the developed learning materials and curricula expect to reflect changes in the market demands.

Timing

- Already happening: learning materials for DS module are provided to deliverables (started on M6 and continues until the latest deliverables on M18).
- Next: collecting a feedback from participants of the learning activities (will happen continuously with releasing feedback analysis results on M24).
- Future: modification of the existent module and development of new modules based on obtained feedback and demand analysis (next releases of updated materials will be on M24 and then on M30)

Technische Universiteit Eindhoven (TU/e)

Partner description

As a university, TU/e's core process is teaching. In the context of the Data Science Center Eindhoven (DSC/e), a new Bachelor program and Master program on data science are launched in 2016 and 2017. DSC/e is the biggest data science initiative in the Netherlands.
TU/e will develop several MOOCs and possibly other courses in EDSA. Using process mining analysis techniques, TU/e and the DSC/e can contribute unique analysis techniques in order to extend current learning analytics capabilities. Within EDSA, DSC/e will apply their process mining techniques on most of the MOOCs developed within EDSA.
TU/e's unique expertise in areas such as process mining, visual analytics and social computing will be disseminated in the EDSA. In the context of the Data Science Center Eindhoven (DSC/e) we also aim to develop new professional education (post academic courses) on data science. Moreover, DSC/e will use EDSA to disseminate its data science toolbox incorporating tools like ProM.

Opportunity: Increased visibility of EDSA data science training

Description

EDSA provides TU/e the opportunity to develop and improve MOOCs on the topic of process mining. This helps both TU/e and DSC/e to increase visibility and reach people otherwise unreachable. The first MOOC on process mining attracted 45,000 registered students, and within all runs so far we have reached over 80,000 students. This order of magnitude of visibility cannot be achieved easily through other means. At the same time, the MOOCs educate many people in the area of process mining, something that would be infeasible otherwise.
Successful online courses part of EDSA can engage large amounts of students, and contribute to the reputation and visibility of TU/e. The success of EDSA benefits TU/e, by showcasing the quality of its teaching through its MOOCs that are part of EDSA.

Impact and relevance

The impact is plentiful. First of all, students and professionals are introduced to TU/e and DSC/e as institutes. Secondly the topic of process mining gains greatly in recognition. This is observable by the fact that many people pro-actively approach us for collaboration, also on the topic of process mining. Furthermore, prospective personnel (at PhD, postdoc and assistant professor level) is attracted to the DSC/e by following the MOOCs.
applying our process mining techniques in the context of learning analytics helps not only the EDSA project and their partners, but also our research since we gain concrete data and related questions to further improve our state-of-the-art techniques.

Commercial aspects

Although hard to measure, the increased market visibility has a positive impact on the number of students for the bachelor and master courses, specifically the ones focussing on data science. Commercially the effect of the MOOCs has been that parties are contacting us pro-actively for contract research, for instance through funded PhD projects. Also other collaborations are initiated through the MOOCs, for instance MSc final projects or case study collaborations. Each of which would be hard to achieve otherwise.

Actions

TU/e plans to keep the MOOCs running and up-to-date, even after the EDSA project has finished, since we see the positive impact it has on a multitude of aspects. In case the upcoming MOOCs from TU/e are very successful in the purpose of this opportunity, TU/e will continue creating new MOOCs to amplify these results.

Timing

The MOOC "Process mining: data science in action" is currently running regular sessions, attracting several thousands of (new) students in each run. The MOOC "Process mining with ProM" launched on July 11, 2016.
In the near future more MOOCs will be delivered by TU/e before the end of the EDSA project; we expect to release one early 2017, and another by the end of 2017.
Beyond the project we expect to keep the MOOCs running, with regular content updates.

EDSA Business Value

This section summarises the expectations given within the previous section but also brings into account the point of view of the different partners in a more focused future-customer perspective.

Expectations of SME partners

ODI

The ODI expects to extract substantial business value both directly from the EDSA's project work as well as subsequent activities:

- Engagement with new audiences and potential clients through consortium relationships, wider network building and increased exposure of ODI's courses and training offering.
- Engaging more with stakeholders in data science, especially through work carried out as part of the demand analysis. This also includes involvement with new industry sectors across the EU.
- Partnering opportunities with the consortium and third parties, for example the EDSA Charter to establish the European Data Science Academy to last beyond the lifetime of the project.
- Development of new products, such as online content and courses, which can be repackaged and sold to other external clients.
- Strengthening business relations with EDSA partners, either existing relationships (e.g. University of Southampton), or new ones (e.g. Open University).

ideXlab

ideXlab contributed to market analysis, collaborated on demand analysis and on the promotion of the curricula and resources developed through open innovation channels. ideXlab expects to extract substantial value both directly from the EDSA's project work as well as subsequent activities:

- The EDSA project commanded adaptations to the ideXlab platform to contribute to the EDSA dashboard.

- Engagement with potential and present clients to present and discuss ideXlab courses and training offering.
- Strengthening partnerships with EDSA project members.

Persontyle

Persontyle contributed to demand analysis using an already established network with industry, academia and leading researchers in the field of data science, machine learning, big data analysis and data engineering. Persontyle worked with EDSA partners on the development of data science curricula and to design, organize and deliver community learning events and training programs. Persontyle expects to extract substantial value both directly from the EDSA's project work as well as subsequent activities:

- Developing a machine learning study guide;
- Develop a new Machine Learning MOOC Program (MOOCs) in collaboration with FutureLearn and OU;
- Professional Certification Program;
- Manage and operate European Data Science Institute / Academy (EDSA Academy) in collaboration with other partners and industry members.

Expectations of research organisations

IAIS

IAIS applied its competences to a quantitative analysis of job offers and multilingual content authoring on the basis of its SlideWiki technology. It collaborated on the delivery of content through face-to-face training.

Further value is expected from further exploitation activities such as enhancing data scientist courses and providing a test bed for innovative data analysis methods.

JSI

JSI contributed the library of thousands of data science related materials at the VideoLectures.NET portal, skills in curricula building for data analytics and natural language processing topics, and data analytics tools and skills in the segment of learning analytics to monitor and visualize the progress of students. JSI expects to extract substantial value both directly from the EDSA's project work as well as subsequent activities:

- Enhancing the operability of VideoLectures.Net portal;
- Extending the PhD and MSc programs at "Jožef Stefan International Postgraduate School";
- Internationalization through multilingual services.

Expectations of academic partners

The OU

The OU applied its competences to the development of the educational curricula of the project and the associated learning resources. It provided its webcasting infrastructure and promoted learning resources via iTunes U. It collaborated on the development of curricula, learning resources and on the delivery of the learning resources.

The OU expects to extract substantial value both directly from the EDSA's project work as well as on increased visibility of the EDSA learning materials.

SOTON

SOTON applied its competences to dissemination and community building through the Web Science Trust and the Web Observatory networks. It collaborated to disseminate training materials to businesses.

More value is expected from further exploitation activities such as:

- Delivering Data Science graduate courses at the University of Southampton;
- Professional data science training.

KTH

KTH used its competences for applying the results of the project in its masters programs as well as in EIT ICT Lab programs where it is involved. KTH collaborated on the development of curricula and on developing and providing training activities.

KTH expects to extract substantial value both directly from the EDSA's project work as well as on quality improvement and distribution of KTH learning material

TU/e

TU/e contributed its expertise on event driven data science (process mining and visualization), the development of analytics-related course materials and the development of ProM-based learning analytics solutions that exploit event data to understand learning processes and performance.

It gained concrete data and related questions to further improve its state-of-the-art techniques and expects to extract additional value by enjoying increased visibility of their research and course material.

Conclusion

The individual exploitation plans vary between the different groups of project partners in EDSA. The academic partners will exploit the results in improving their educational material and distribution channels; the research partners will also improve their courses but also test and deploy their research outcomes in the context of learning analytics and learning delivery. The commercial partners will exploit adjacent business opportunities related to data science learning.

These individual exploitation plans will be complemented by the main joined exploitation activity, i.e. the EDSA Charter.