



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

D2.7 Lessons learned and best practices from the production of learning resources

Deliverable Editor: **Alexander Mikroyannidis (OU)**
Deliverable Reviewers: **Misha Matskin (KTH), Huw Fryer (SOTON)**
Deliverable due date: **31/01/2018**
Submission date: **26/01/2018**
Distribution level: **P**
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	12/01/2018	Alexander Mikroyannidis	Version for internal review.
0.2	22/01/2018	Alexander Mikroyannidis	Revised version.
1.0	26/01/2018	Alexander Mikroyannidis	Final QA.

Table of Contents

Change Log.....	2
Table of Contents.....	3
List of Figures.....	3
1. Executive Summary.....	4
2. Introduction.....	5
3. Lessons learned from applying the EDSA values.....	8
3.1 Lessons learned from applying the EDSA curriculum design values.....	8
3.2 Lessons learned from applying the EDSA curriculum delivery values.....	9
4. Conclusion.....	11

List of Figures

Figure 1: The EDSA production process for curricula and courseware.-----	5
Figure 2: Channels for the delivery of EDSA learning materials - The EDSA courses portal, VideoLectures and FutureLearn.-----	6

1. Executive Summary

This deliverable reports the overall lessons learned and best practices acquired from the work conducted within WP2. The main objective of WP2 has originally been the production of the curricula and learning resources of the project in multiple formats, languages and for a variety of learning purposes and learning contexts. Based on the demand analysis conducted in WP1, as well as the evaluations taking place in WP3, a core curriculum has been developed and refined in order to target the real needs of data scientists and the European data industry. The developed curriculum has been supported by learning resources delivered via a variety of pedagogical methods and platforms.

As previously reported in D2.5 (M24) and as a result of the M18 project review, the focus of the project has been shifted towards addressing the supply of training materials in order to bridge the data science skills gap. As a result, WP2 has extended the EDSA courses portfolio to include a wider range of courses offered by renowned educational institutions both inside and outside the project consortium. These courses have been selected based on their relevance to the EDSA curriculum and the EDSA demand analysis.



2. Introduction

The main objective of WP2 was originally the production of curricula and learning resources driven by the demand analysis conducted in WP1. In order to address the demand for data science skills, a participatory approach was initially adopted by WP2 for the design and production of bespoke curricula and courseware, as shown in Figure 1.

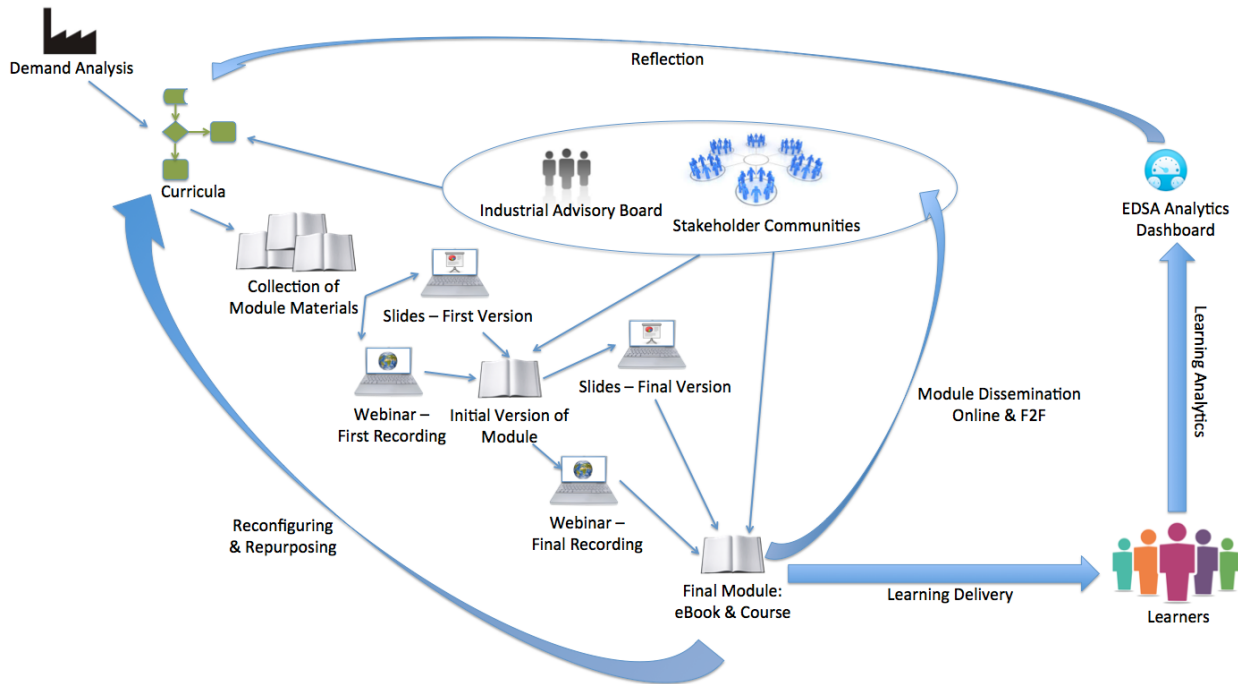


Figure 1: The EDSA production process for curricula and courseware.

WP1 has been monitoring trends across Europe in order to assess the demands for particular data science skills and expertise, using automated tools for the extraction of data science job posts, as well as interviews with data science practitioners. WP1 also established an Industrial Advisory Board representing a mix of sectors to ensure that project activities continue to meet changes in the demands on data science across Europe.

Starting from the results of this demand analysis and input from the Industrial Advisory Board, WP2 has created relevant data science curricula to meet the outlined training needs. A multidisciplinary course writing team has been developing in parallel a repository of relevant source materials, draft modules to be placed online, as well as materials for webinars. The draft modules have then been iteratively revised based on the feedback received from the Industrial Advisory Board, from the face-to-face training activities, as well as from monitoring the main communication channels used by the communities of stakeholders. The analysed feedback has been used to restructure, finalise and publish the module content via different educational platforms, including the EDSA courses portal,¹

¹ <http://courses.edsa-project.eu>

VideoLectures,² as well as platforms of Massive Open Online Courses (MOOCs) including FutureLearn³ and Coursera⁴ (see Figure 2).

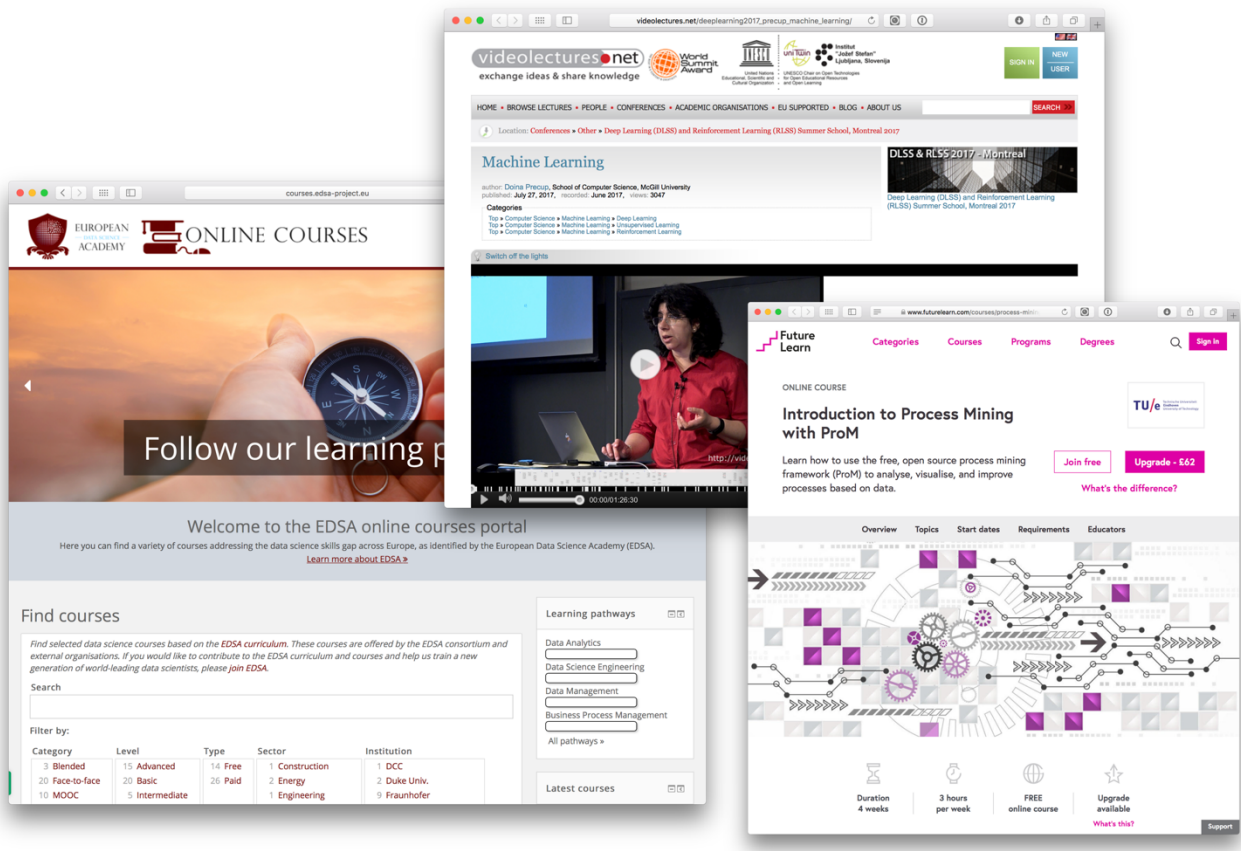


Figure 2: Channels for the delivery of EDSA learning materials - The EDSA courses portal, VideoLectures and FutureLearn.

Additionally, Learning Analytics were incorporated into our online delivery, allowing us to collect data related to the learning experiences of our users, which offered feedback into our curricula design. Based upon the Learning Analytics data and the feedback from our stakeholders, we reconfigured and repurposed modules for different learning contexts, thus initiating new cycles of the production process.

As a result of the project's M18 review, the focus of WP2 was shifted from the production of learning resources to the curation of a courses portfolio aggregating a wider range of high quality learning resources, either offered by project partners or by third parties. This shift of focus aimed at closing the gap between the demand of data science skills across Europe and the supply of learning materials suited for offering the required skills to job seekers.

The EDSA courses portfolio has thus been extended to include additional courses offered by renowned institutions both inside and outside the project consortium. These courses are available as:

² <http://videolectures.net>

³ <https://www.futurelearn.com>

⁴ <https://www.coursera.org>



- *Massive Open Online Courses (MOOCs)*: These are online courses aimed at unlimited participation and open access on the web. They are available on external MOOC platforms, including FutureLearn and Coursera.
- *Face-to-face courses*: These courses are taught face-to-face. Face-to-face learning (or in-person learning) is any form of instructional interaction that occurs “in person” and in real time between teachers and students or among colleagues and peers.
- *Online courses*: These courses are taught online via Learning Management Systems (LMSs) like Moodle or Sakai. A subset of these courses consists of self-study learning materials available as Open Educational Resources (OERs), which learners can study at their own pace, as there is no predetermined start or end date.
- *Blended courses*: These courses are taught in a blended way (face-to-face and online). Blended learning is a formal education program in which a student learns at least in part through delivery of content and instruction via digital and online media with some element of student control over time, place, path, or pace.

The main criteria for the selection of both internal and external courses for inclusion in the EDSA courses portfolio have been the *EDSA curriculum* and the *EDSA demand analysis*. Courses have been selected based on their potential of addressing the EDSA curriculum topics as well as the training needs of data scientists as identified by the EDSA demand analysis. With regards to compliance with the EDSA demand analysis, courses have been evaluated against the recommendations of the Study Evaluation Report (D1.4). With regards to compliance with the EDSA curriculum, courses have been evaluated against the latest version of the curriculum and the topics it addresses. We have also established a process for monitoring changes to the demand analysis and aligning our curriculum accordingly. This process is documented in Data Science Curricula 3 (D2.3).

Linking the demand for data science skills with the supply of learning resources that offer these skills is crucial for bridging the data science skills gap. Towards this goal, EDSA has developed an interactive dashboard⁵ that enables its users to explore both the current data science skills demand and supply. The EDSA dashboard enables its users to explore both the current data science skills demand and supply. Users of this dashboard are able not only to explore the current demand in the data science market, but also find learning materials and training relevant to the skills they will need to secure a specific job position. Additionally, users are supported in building personalised learning pathways, consisting of courses and learning materials that will help them reach their learning goals. The EDSA dashboard is presented in more detail in the Demand and supply analysis report (D1.5).

The remainder of this deliverable reflects on the lessons learned and best practices gained from the WP2 work. In particular, we revisit the EDSA values, which were established early in the project, and we reflect upon these values in light of the lessons learned throughout the duration of the project.

⁵ <http://edsa-project.eu/resources/dashboard/>

3. Lessons learned from applying the EDSA values

The EDSA values⁶ were established within the first few months of the project in order to drive the development and delivery of the project's curricula and learning resources. The EDSA values have been grounded on the pedagogical and technological expertise of project partners, as well as on the lessons learned from the previous involvement of EDSA partners in relevant research projects, such as EUCLID⁷ and LinkedUp.⁸

The EDSA values define the principles that the EDSA curriculum and learning resources should adhere to. The EDSA values also comprise a set of best practices that can be used as guidelines generally for the development of online courseware.

In the following sections, we revisit the EDSA values and reflect on how the WP2 work was aligned with these values, as well as the lessons learned.

3.1 Lessons learned from applying the EDSA curriculum design values

- **Industry Aligned:** The EDSA curriculum has been designed in accordance with the expectations of EU industrial sectors connected to data science, providing industry-standard scenarios and tools. In particular, the Industrial Advisory Board established by the project has contributed to this value in great lengths, by offering feedback and reshaping the project's curricula and learning resources. This has allowed our learning resources to reach a wider audience and target specific industry sectors.
- **Industry Standard Tools:** Our compilation of open source data science tools has offered learners experience with tools customary to the industry and their specific sector. In particular, our learning resources have focused on widely accepted tools with a strong user base within the industry. In this way, the produced learning resources have become relevant to the everyday practices of data scientists.
- **Real Data:** Learners utilising the EDSA curriculum have access to a number of large-scale open datasets to perform their learned data science skills, enabling real-world data science on real-world data. This is especially true for MOOCs and online courses either produced or endorsed by EDSA that utilise a wide range of datasets with real applications on real usage scenarios. Like before, this approach has allowed our learning materials to appeal to a wider industry audience as they integrate seamlessly with their everyday practices.
- **Open Design:** The EDSA curriculum has been designed from user, research, industry and professional recommendations and feedback has been taken into account from all across the EU, ensuring that the curriculum meets the needs of the industry, academia and the wider market. As mentioned before, the project's Industrial Advisory Board has brought together representatives from key industrial and academic sectors across Europe and beyond. This has allowed us to collect rich feedback from representatives of the world-wide industrial and academic data science community, in order to contribute to the training of a new generation of world-leading data scientists.
- **Expert Provision:** The EDSA curriculum has been designed by world-class professional and academic experts in data science. In particular, we have employed experts in different data science areas for the development and delivery of our curricula and learning resources. In the cases that the required expertise was not available within the project consortium, we have identified and endorsed appropriate learning materials of high quality that are offered by

⁶ <http://edsa-project.eu/overview/edsa-values/>

⁷ <http://www.euclid-project.eu>

⁸ <https://linkedup-project.eu>



organisations outside of the project consortium. In this way, the project's course portfolio has been extended to include high quality learning resources developed both inside and outside of the project.

- **Modular:** The EDSA curriculum is flexible and adaptable to educator requirements and the needs of their learners. In order to ensure that this is true, we have performed our demand analysis across a variety of industry sectors. In this way, we have aimed to deliver learning resources that address a wide range of learning requirements, as well as resources that can be reused and repurposed by releasing them as Open Educational Resources under Creative Commons licenses.
- **Transferrable:** Skills learned through the curriculum can be utilised across a range of data science roles, occupations and countries throughout the EU. As mentioned before, the project has dedicated efforts to performing an extensive demand analysis across Europe. This analysis has identified the data science skills that are currently mostly in demand by the industry throughout Europe and beyond. As a result, the EDSA curriculum and the associated learning resources have been developed around these skills.
- **Concise Learning Goals:** All EDSA courses have been aligned with clear learning goals depicted by a specific aspect of the data science role. This has allowed learners to identify the courses that best suit their learning needs and preferences. In addition, learning pathways have been provided to enable learners to navigate through the content, selecting what is useful to them. Learners may follow these learning pathways, or further customise and personalise them, as well as reflect on their learning progress.
- **Addressing the Whole Data Value Chain:** Data scientists are made aware of the techniques and stages of the whole data science value chain through the use of easily understandable narratives. This is especially true in the free and open online courses offered by ODI, such as the module "Finding Stories in Data",⁹ where clear and concise narratives and storytelling drive the whole learning experience.

3.2 Lessons learned from applying the EDSA curriculum delivery values

- **Multilingual:** The EDSA courses have been delivered across a number of European languages to extend their reach and enable others to use the EDSA curriculum. This value has been followed to a certain extent, as it requires additional resources dedicated to the translation and customisation of learning materials for different countries and learning contexts. It became apparent quite early in the project, that the resources available within the EDSA consortium would allow primarily for the production of English language resources. However, the consortium has published a substantial part of its learning materials as Open Educational Resources, thus allowing them to be translated and repurposed by the community.
- **Multimodal:** The EDSA courses have been provided in a number of modes to suit skill levels and format preferences, such as MOOCs, eBooks and slide decks. As mentioned in the Introduction section of this deliverable, the EDSA courses portfolio spans across a variety of pedagogical models and employs different delivery channels and formats in order to address different learning contexts and audiences. The EDSA courses also cover all types of pedagogies, from the traditional face-to-face pedagogical model, to the more recent trends in online education (MOOCs and blended learning).
- **Multi-Platform:** EDSA has utilised a wide range of platforms in order to remain accessible and available to a large body of data science learners. We have primarily used VideoLectures and FutureLearn, which is the largest European MOOC platform founded by The Open University,

⁹ <http://courses.edsa-project.eu/course/view.php?id=52>

in order to maximise outreach and uptake of the EDSA learning materials. As a result, we have appealed to thousands of learners worldwide that have enrolled and attended the EDSA MOOCs.

- **Cutting-Edge Quality:** The EDSA learning materials have been subject to a series of design iterations that encapsulate the latest research and professional practice, prior to their launch. This has been true for the materials developed within the project, which have followed the iterative production process described in the Introduction section of this deliverable. Additionally, external learning resources have been evaluated before being incorporated into the EDSA courses portfolio, based on their potential of addressing the EDSA curriculum topics as well as the training needs of data scientists as identified by the EDSA demand analysis.
- **Reflective and Quantified:** The EDSA learning materials have been delivered with data and analytics in mind, providing learners with quantified measures and analytics to reflect on their aptitude, skills and strengths. Learning Analytics have been incorporated into the online delivery of EDSA courses, allowing the collection of data related to the learning experiences of learners. Additionally, self-assessment exercises and quizzes have been incorporated in the EDSA courses, thus enabling learners to monitor and reflect on their learning progress.
- **Hands-On:** The EDSA course materials have been delivered in a way to emphasise a constructivist hands-on approach, meaningfully applying knowledge to real tools and data. As mentioned before, the EDSA learning resources have focused on widely accepted tools with a strong user base within the industry. In addition, the courses either produced or endorsed by EDSA have utilised a wide range of datasets with real applications on real usage scenarios, thus appealing to a wider industrial audience.



4. Conclusion

This deliverable has reported the overall lessons learned and best practices acquired from the work conducted within WP2. In particular, we have revisited and reflected upon the EDSA values for the design and delivery of curricula and courseware. The EDSA values comprise a set of guidelines and best practices that can be reused across different contexts related to designing and delivering online courseware. The EDSA values can be therefore regarded as a sustainable outcome of the project, having the potential to inform future initiatives relevant to the production and delivery of online courseware for data science or other data-related fields.

In retrospect, designing curricula and courseware for data science has been an inherently difficult task facing a number of challenges, most notably the speed at which this field is changing. Increasing amounts of data lead to challenges around data storage and processing, not to mention increasing complexity in finding the useful story from that data. New computing technologies rapidly lead to others becoming obsolete. New tools are developed which change the data science landscape. These all occur at such a rapid pace that teaching data science requires an agile and adaptive approach that can respond to these changes. In the context of EDSA, we have carried out revisions guided by the EDSA values to the curriculum and the associated learning resources throughout the duration of the project, in order to reflect the most up-to-date needs of the data science community.