# EDISON

## Education for Data Intensive Science to Open New science frontiers

**Project no.** 675419
**Coordination and Support Action**
**Funded by the Horizon 2020 Framework Programme of the European Union**

| | |
|---|---|
| **Call identifier:** | **H2020-ICT-2015-1** |
| **Topic:** | **INFRASUPP-4-2015 - New professions and skills for e-infrastructures** |
| **Start date of project:** | **1 September 2015 (24 months duration)** |


# Deliverable D3.3

# Final Report on the Use Cases support design

| | |
|---|---|
| **Due date:** | 31/08/2017 |
| **Submission date:** | 31/08/2017 |
| **Deliverable leader:** | FTK |

Dissemination Level

| | | |
|---|---|---|
| ☒ | PU: | Public |
| ☐ | PP: | Restricted to other programme participants (including the Commission Services) |
| ☐ | RE: | Restricted to a group specified by the consortium (including the Commission Services) |
| ☐ | CO: | Confidential, only for members of the consortium (including the Commission Services) |

EDISON project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 675419

# List of Contributors

| Participant | Short Name | Contributor |
|---|---|---|
| Research Institute for Telecommunication and Cooperation | FTK | Jana Becker, Dominic Heutelbeck, Matthias Hemmje |
| Engineering Ingegneria Informatica SPA | ENG | Andrea Manieri |
| University of Amsterdam | UvA | Adam Belloum, Yuri Demchenko |
| University of Stavanger | UiS | Tomasz Wiktorski |

# Change history

| Version | Date | Partners | Description/Comments |
|---|---|---|---|
| 0.1 | 26/06/2017 | FTK | Initial draft and structure of deliverable |
| 0.2 | 17/07/2017 | UvA / FTK | Chapter on Accreditation and further input |
| 0.3 | 03/08/2017 | FTK | Introduction part, Piloting Use Cases A and B supplemented, screenshots added, Annex C expanded (more material to come) |
| 0.4 | 08/08/2017 | FTK | Further additions |
| 0.4 | 13/08/2017 | FTK | Summary and further refinement |
| 1.0 | 30/08/2017 | FTK | Cleanup, Taking up quality assurance comments, and integration of final contributions |

## Executive summary

The work package WP3 - Development and Reference Implementation Strategy had the goal to connect the theoretical results achieved in work package WP2 into concrete actions resp. use cases for the target communities. This has been achieved in support of the piloting use cases by involving different stakeholders related to Data Science and corresponding Accreditation and Certification.

The successful engagement and process of EDISON Data Science Framework (EDSF) validation and implementation is described, following an iterative approach along the four components Data Science Competence Framework (CF-DS), Body of Knowledge (DS-BoK), Data Science Model Curriculum (MC-DS), and the Data Science Professional Profiles (DSPP).

The feedback regarding the EDSF was positive and the structured guideline to use the theoretical framework could be applied smoothly. EDSF was recognized by the target communities and stakeholders to ensure usability and quality for the design and development of new education and training offerings in the important branch of Data Science.

The engagement events, like the Champions Conference and consequently hands on training and bilateral discussions were very productive. The workshops and talks were perceived as informative and future-driven, but especially the encompassing exchange across different implementations of the Data Science curriculum for different target professional groups was confirmed as valuable for all involved participants.

The EDISON team is very proud to pronounce that the Champions Conference will get into continuity and will be established in the Data Science landscape as a regular event for scientific and industry-driven communities.

TABLE OF CONTENTS

## List of figures

# List of tables

# 1 Introduction

This deliverable describes the final version of the *Piloting Use Cases* support design for the Use Cases A, B and C (*Tasks 3.2, 3.3*).
The final version of the *EDISON Data Science Model Curriculum* (*Task 3.1*) was published in Deliverable D2.3 [3].
This approach has been agreed in consultation with the Project officer, Amendment 675419-14.
The final outcomes of *EDISON Online Education Offering (Task 3.4)* will be described in Deliverable D3.4.

Data Science is an emerging field of science, which requires a multi-disciplinary approach and should be built with the strong link to Big Data and data driven technologies that created transformational effects to all research and industry domains, and consequently require re-thinking and re-design of both traditional educational models and existing courses. However, at present time most of the existing university curricula and training programs are built based on available courses and cover only a limited set of academic subjects related to a full Data Science Body of Knowledge covering only a limited set of Knowledge Areas and Professional Profiles as defined in the project. This potentially may create gaps in knowledge and competences of the future Data Scientist graduates for their smooth integration in the real working environment (in both industry and academia).

The EDISON consortium therefore developed the EDISON Data Science Framework (EDSF) as a set of instruments to support different stakeholder groups in the establishment of the Data Scientist profession and match the market requirements.

The Piloting Use Cases evaluated the EDISON Data Science Framework and exemplify the adoption in a number of practical settings and the respective associated stakeholder groups. Use Case A addresses the academic education of Data Scientists within dedicated university programs, whereas Use Case B focusses on re-skilling/change of career path as part of lifelong learning/continuing professional education. Finally, Use Case C investigates into European-wide harmonization of Data Science accreditation.

## 1.1 Purpose

EDISON developed a new *Data Science Framework (EDSF)* to support Universities and Scholars with a *Body of Knowledge (DS-BoK)* implemented by a *Data Science Model Curriculum (MC-DS)* as target competence profiles to structure their curricula as well as their corresponding educational offering, i.e., study programs and courses. Furthermore industries, training and education providers should be supported with the *Data Science Competence Framework (CF-DS)* and the *Data Science Professional Profiles (DSPP)* to define employment policies and to define and execute life-long learning activities.

The design of the Model Curriculum has taken an accommodation of specific requirements related to training and education corresponding to stakeholders across scientific domains, Research Infrastructures, industry, and the public sector into account.

The Piloting Use Cases were conceived to support the validation and iterative refinement of the proposed CF-DS, DS-BoK, MC-DS, and DSP taxonomy through diverse learning paths and across learning models (e.g., residential, online, blended) through the engagement of stakeholders for Data Science education and training, but also the accreditation of corresponding Data Science education and training programs. The Piloting Use Cases and the champions engaged therein have served the purpose of validation and fine-tuning EDISON outcomes before exposure to larger communities.

Furthermore, the inclusion of the champions and stakeholders has led to the alignment of the results with the needs of the target groups and on the other hand a great willingness to adapt the results and to incorporate them into future training and education offerings.

## 1.2 Scope

Work Package 3 was responsible to define and develop a Model Curriculum for Data Science and to support a pilot implementation of the EDSF instruments along the described Use Cases.
Deliverable D3.1 pictured the initial version of the MC-DS and specified the Piloting Use Cases.

The presented deliverable D3.3 documents the implementation of Piloting Use Cases and accumulates the input and feedback received from the champions and wider community discussion regarding the EDSF and the final legacy of the EDISON project.

As stated in the Description of Action [1], EDISON does not aim to develop any concrete residential or online modules/units. However, the availability of the CF-DS, the DS-BoK and the experience after the implementation of MC-DS by a number of piloting Use Cases will reduce the complexity of developing new modules/courses and programs at different universities engaged in EDISON and cooperating with the project.

## 1.3 Structure of Deliverable

The Deliverable follows the Deliverable 3.1 and contains the summary of Piloting Use Cases A, B and C.
Chapter 2 is describing the Use Cases A and B, including the subsections on Use Case Support Design, the Implementation Strategy, the description on EDSF Application by several Champions and the Implementation of Online Training Courses into the EDISON Online Education Platform. The further engagement regarding the Champions Conference and the Online Content is outlined afterwards.
Chapter 3 is characterizing the Use Case C on Accreditation. Chapter 4 is summarizing the results and describing the outlook on future activities regarding the EDISON legacy.

# 2 Piloting Use Cases (A and B)

Universities are claiming the role to educate (young) students to understand and develop their unique capabilities and careers. On the other hand, Life Long Learning Training is focusing on re-skilling or changing the career path. To include both perspectives, EDSION was working together with experts from the academic education sector and from continuous professional training in Life Long Learning scenarios to demonstrate the suitability of the overall proposal.

An “EDISON Internal Champion” is an institution that wants to align its Data Science program (either running or planned) as early-adopters with EDISON Data Science Framework components. The project and its champions collaborate to develop and share relevant information and other resources needed, and to share results of such endeavor which contribute to better fulfill their respective institutional mandates.
The corresponding Terms of Reference defining roles/responsibilities of Internal Champions and the corresponding benefits from their participation in the project are described in Appendix A. Internal Champions’ Terms of Reference.

The Champions participated to assess, align, and adopt EDISON Data Science Framework components and therefore support the design and development of Data Science Education and Training content during EDISON and, especially in the long term, after the runtime of the project.

That has been achieved by pilot training sessions, the performance of face-to-face and online education and training opportunities and individual discussion sessions for the different areas of target groups.
The targets were:

- to explain the ideas behind the EDSF components, like the CF-DS, MC-DS, Learning Outcomes (LO) und Learning Units (LU), DS-Bok and the DSPP and corresponding taxonomy,
- to give guidance how to apply the components,
- to support a first mapping of existing courses and content and
- to give advice for the future adaption and application of the EDISON outcomes into further developments of new Data Science offerings.

## 2.1 Use Cases Support Design

The development of Data Science Education and Training courses in EDISON is supported by two distinct (but non-orthogonal) piloting use cases:

- **UC A University Champions** – "Universities and Scholars will use EDISON outcomes as target competence profiles to structure their curricula (based on the MC-DS) as well as their corresponding educational offering, i.e., Data Science programmes and associated modules/units.“

- **UC B Lifelong Learning/Training** – "Industries will use EDISON outcomes to define employment policies and of course by trainers and education institutions to define and execute lifelong learning activities. “

Hence, UC A targets the application of EDISON outcomes for supporting academic education (“to educate young students to understand and develop their unique capabilities and careers”). UC B addresses Continuous Professional Education/Lifelong Learning scenarios (“re-skilling or planning the career path”).

The implementation of both above-mentioned Use Cases requires the active participation of “EDISON Internal Champions”, which should represent different types of educational profiles as well as address a diverse range of target sectors (such as science, public institutions, industry).

### 2.1.1 Appraisal and Selection of Internal Champions

Academic partners UvA and UiS were natural candidates for being EDISON Internal Champions. However, the engagement of the principal actors in the relevant Data Science programs (which are not necessarily involved in EDISON) needed to be secured.

For non-academic partners ENG and FTK the appraisal of eligible candidates drew on the exploitation of pre-existing working relationships/networks and/or personal contacts. Nominations were circulated through the project mailing list for discussion.

The actual selection process was determined by multiple factors/multi-faceted criteria (such as coverage of MSc and BSc programs, planned vs running DS program, geographic location, addressing Continuous Professional Education issues, but also resource constraints associated with candidates and/or their mentors). Finally, the list of EDISON Internal Champions was then approved/selected by all WP3 partners.

The following EDISON Internal Champions have been appraised and selected:

- The Department of Computer Science at the **University of Bedfordshire** is offering a full 3-year Bachelor degree BSc (Hons) Data Science programme. The programme is offered in the current academic year 2015-16 [11]. The programme has been designed as being application-oriented adhering to market demands. It benefits from a strong collaboration with the MK:Smart project in Milton Keynes (http://www.mksmart.org) that allows students to work on real-life Data Science problems with real data, adhering to demands coming from industry and research. The Data Science programme is delivered by members of a Data Science research group within the Department of Computer Science whose expertise comprise areas such as data analytics, data mining, machine learning, information engineering and retrieval, cybersecurity, multimedia, visualisation, digital libraries and databases.
- The Department of Engineering at the **University of Perugia** is offering a Data Science programme [12] organized in two levels from the academic year 2016-2017. The first level consists of a two-year graduate course in Computer Engineering, with a course curriculum in Data Science. This course can be accessed by student having an undergraduate degree in Electronic Engineering or Computer Engineering. It includes classes focused on Data Science, such as big data analytics, big data management, data mining, machine learning, virtual networks and cloud computing, cybersecurity, and visualisation. The second level consists of a post-graduate international, industry driven, Master Certificate in Data Science. It is organized in collaboration with enterprises, and has the purpose of specializing the competences of students in strategic business sectors.
- The Institute of Computer Science ("Informatik") at the **Goethe University Frankfurt** is currently evaluating the possibility to offer a Master in Computer Science with focus on Data Science. This program will be done in cooperation with the Institute of Mathematics of the Goethe University. The two institutes are part of the Department of Computer Science and Mathematics (FB12 [13]). Currently our both institutes already offered a variety of courses/seminars/labs related to Data Science [14], such as Machine Learning, Statistics, Big Data Stores, etc. If the evaluation is positive the new Master in Computer Science with focus on Data Science will be introduced by the end of 2017. The coordinator for this evaluation is Prof. Roberto V. Zicari.
- The **University of Amsterdam** is member of the Amsterdam Data Science initiative [15], which aim at developing data science talents and technology within the Amsterdam region. For the 2016-2017 academic year the University of Amsterdam in collaboration with Free University Amsterdam is offering a number of Data Science oriented programs both at MSc, and postgraduate levels [16]. At MSc level, a new specialization within the Computer Science program “Big Data Engineering” is focusing on the technologies based on which Big Data infrastructures are built. The Big Data Engineering specialization [17] covers fundamental knowledge addressed in 5 core courses namely: Large-Scale Data Engineering, Web Data Processing Systems, Web Services and Cloud-based Systems, Information Visualization, and Data Mining. At postgraduate level, the University of Amsterdam in collaboration with the Dutch National supercomputing center (SURFsara) offers a series of lectures and workshops targeting a wide range of potential big data users helping them to Enhance the accessibility of the HPC (High Performance Computing) facilities to solve Data and CPU intensive applications [18].
- The Department of Electrical and Computer Engineering at the **University of Stavanger** has been offering a course in Data Intensive Systems [19] since 2013, it also has courses in Machine Learning, Distributed Systems, etc. It is now planning to build a general Master program in Data Science basing on its own competences in collaboration with Department of Mathematics and Natural Sciences.
- University of Stavanger is a partner with **Lodz University of Technology** in the European Consortium of Innovative Universities. The Department of Medical Electronics at the Lodz University of Technology has a bachelor program in Bioengineering [20], which is a Data Science program with domain

application in bioengineering. Through the project general improvements are planned to the program based on Model Curricula. Synergies will also be explored between competences at the University of Stavanger and Lodz University of Technology.

- The **Engineering Training School "Enrico della Valle"** [21] is located in Ferentino just outside Rome, where Engineering runs its methodological, technological and managerial training courses. The school is organised as a post-university campus. Several training programs have been set up focusing on optimization and development of vertical skills and professional development for key resources based on: lessons in the classroom, work on projects alongside the professionals, project management, virtual classes, self-study, case studies and communications. The School is recognized as a Certifying Body of international associations and market players (i.e. PMI, ITIL, Prometric, COBIT, Prince2, Pearson). The Academy's range of courses is completed by e-learning services: a catalog of courses to be completed remotely from one's own computer or on Engineering's technological platform, benefitting from tools and resources for tutoring, monitoring, tracking of use and verification of outgoing knowledge.
- The **Persontyle School of Data Science** [22] is an education initiative to help meet the world's demand for professionals and leaders skilled in developing and utilizing automated and intelligent methods of using data as a strategic resource. They aim to bring accessible, affordable, practical, and interactive data science and engineering education to the world. They offer training programmes [23] and tailored corporate learning solutions to cover the concepts, technology and applied practices which are needed throughout the entire lifecycle, from asking the relevant questions to making predictions using machine learning models and visualizing results. Whether someone is an established professional or new to the fields of data science, data engineering or machine learning, the training shall provide the expertise you need to make sense of the data.
- The **Lucerne School of Information Technology** [24] is located in Zug-Rotkreuz between Lucerne and Zürich in Switzerland. This school is a new department of the Lucerne University of Applied Sciences and Arts [25] that starts its operation in September 2016, resulting from an integration of existing degree programmes and research units at the School of Business and the School of Engineering into a new, independent university school. The Lucerne University of Applied Sciences and Arts offers BSc and MSc degree programmes, continuous professional education (CAS, DAS, MAS), and conducts research on national and international level. A new major in Data Engineering and Data Science [26] for BSc in information technology and BSc in business informatics will start in September 2016.
- The **Jožef Stefan Institute** (JSI) [27] is the leading Slovenian scientific research institute, covering a broad spectrum of basic and applied research. The subjects concern production and control technologies, communication and computer technologies, knowledge technologies, biotechnologies, new materials, environmental technologies, nanotechnologies, and nuclear engineering. The mission of the Jožef Stefan Institute is the accumulation - and dissemination - of knowledge at the frontiers of natural science and technology to the benefit of society at large through the pursuit of education, learning, research, and development of high technology at the highest international levels of excellence. JSI is leading the EU funded project "European Data Science Academy (EDSA)".

### 2.1.2 Use Case Characterizations

The following tables depict brief characterizations of the respective piloting use cases of types A, B to be implemented within Task 3.2. Furthermore, the assignments of EDISON mentors to EDISON Internal Champions are stated.

**Table 1: UC A University Champions**

| **UC A (Internal) University Champions** | | |
|---|---|---|
| *Title* | *Internal Champion* | *Mentor* |
| Verification of BSc DS programme, supporting the approval of future (derived) programmes. | University of Bedfordshire | FTK |
| Refining the post-graduate Master Certificate consistent with EDISON Data Science Framework components, promotion of (student-level, between universities) collaboration. | University of Perugia | ENG |
| Endorsing the establishment of a MSc in Computer Science with Data Science focus, validation of curriculum w.r.t. EDISON Data Science Framework components, identification of gaps | Goethe University Frankfurt | ENG |
| Design of BSc and MSc curriculum for Data Science using existing | University of Stavanger | UiS |

| | | |
|---|---|---|
| courses, promoting Data Intensive Systems as integral part of Data Science curricula. | | |
| Update of BSc Biomedical Engineering according to the EDISON Data Science Framework, providing domain-specific instance thereof. | Technical University of Lodz | UiS |
| Establishment of Data Science curriculum focusing on Data Management, accommodation of Data Science profiles within Amsterdam Data Science. | University of Amsterdam | UvA |
| Verification and alignment of consecutive BSc programmes with major in data engineering and data science which start in September 2016. | Lucerne School of Information Technology | FTK |

**Table 2: UC B Lifelong Learning/Training**

| **UC B Lifelong Learning/Training** | | |
|---|---|---|
| *Title* | *Internal Champion* | *Mentor* |
| Alignment of learning programmes with EDISON Data Science Framework components and design of online courses for Lifelong Learning/Training | The School of Data Science | FTK |
| Designing online courses for Lifelong Learning/Training and re-skilling in the areas of Data Science and Computing for Data Scientists. | Engineering Training Center | ENG |
| Possible development of new MAS (Master of Advanced Studies) professional education in data science building on the EDISON Data Science Framework. | Lucerne School of Information Technology | FTK |
| Designing online courses for Lifelong Learning/Training and re-skilling in the areas of Data Science and Computing for Data Scientists. | Jožef Stefan Institute | FTK |

## 2.2 Internal Champions Implementation Strategy

The implementation strategy for the Piloting Use Cases A, B adopts a phased approach:

1. The first phase (Phase I) focuses on the engagement of University Champions (UC A), asking them to reflect on and apply EDISON Data Science Framework components in the context of their respective (planned) Data Science programmes.
2. Building on initial experiences from Phase I, the second phase (Phase II) then broadens the scope to the engagement of professional stakeholders from Lifelong Learning/Training (UC B), aiming at the alignment and contextualization of their programs with EDISON Data Science Framework components. This part was carried out in a more discussion oriented way and a bilateral process between mentor and non-academic partner.

### 2.2.1 Methodology

The EDISON Piloting Use Cases are conceived as *agile*, i.e. their implementation was iterative and evolutionary. The Internal Champions are the *Primary Actors* (their mentors are supporting *Secondary Actors*), whereas the EDISON Data Science Framework characterized the System under discussion (see, e.g., [42] for a brief description of use cases). The Scenarios are than defined by engaging the Primary Actors in the application of the components of the EDISON Data Science Framework (EDSF). They reflected based on their individual requirements regarding Data Science education and training curricula, validated the usability of the provided EDSF along a structured approach and execute the implementation of the system as far as possible. The corresponding components of EDSF are:

- **Data Science Competence Framework** (CF-DS) [3, 4]) – Defines and classifies competence groups essential for Data Science profession that should supported by academic and training programs.
- **Data Science Body of Knowledge** (DS-BoK) [3, 5], defines updated version is provided in Appendix C) – Determine coverage of knowledge areas/topics in offerings.
- **Data Science Model Curriculum** (MC-DS) [3, 6]) – Mapping of courses to Data Science Model Curriculum (including specification of eligible paths), contextualization of programmes with professional profiles, elicitation of/reflection on eventual gaps and assessment of the expected impact of further offerings.

- **Data Science Professional profiles** (DSP Profiles) [3, 7]) – Determines professional profiles targeted in offerings.
- **Education and Training Resources Inventory** (D2.2 Chapter 2 [2], and online directory [28]) – Validate programme information, correct/supplement information (exchange format). Provide course information for the online directory (exchange format).

All components of the EDISON Data Science Framework are published in the EDISON website Library [29] as Open Source documents CC BY license that are regularly updated based on the community feedback and in particular feedback from Champions. The final version (Release 2) was published on 3 July 2017.

The Use Cases were following the approach of:

- Engagement
- Validation
- Implementation
- Continuation

The engagement was ensured by events and bilateral exchange (cf. 2.2.2), validation and implementation were performed by an illustrated walkthrough and the so-called EDSF application (cf. 2.3) and the implementation of online training courses (cf. 2.4). The continuation of the initial work with the target communities will be progressed by future events and the population of the online platform (cf. 2.5).

The individual Use Case scenarios are initialized by a reflection (desk-based, assessing scientific and technical soundness, scope and relevance, readability, appearance and structure, usability, etc.) of the particular EDISON Framework constituent, relying on the documents provided by the project. Then it is followed by hands-on application w.r.t. the Internal Champions own offerings, supported by the respective Mentors. Experiences, issues, and insights were captured systematically and reported back to the EDSF development team for consideration and management/implementation of changes, if need be.

### 2.2.2 Modes of Engagement

Most interactions were bilateral (either remote or Face-to-Face), between Internal Champion and its associated Mentor. However, dedicated **Pilot Trainings** about the "*EDISON Data Science Framework in use*" were organized in collaboration with WP5 as part of the 3 *Use Case Events* foreseen by the Description of Action (cf. [1]). To capture the Champions validation and mapping of EDSF components to their own offerings, a so-called Illustrated Walkthrough document and a corresponding excel spreadsheet were developed and applied.

Dedicated Face-to-Face meetings were arranged between UC B Internal Champions and their mentors for the discussion of addressed educational target groups, general assessments of the applicability of EDISON Data Science Framework components and the elicitation of expected outcomes.

### 2.2.3 Pilot Trainings

Following the initial planning, the following Pilot Trainings were conducted:

- 1st EDISON Champions Conference[1] (July 13-14, 2016, Brockenhurst, UK) - Session II: "*Learning from the Champions: presentations from pioneering early adopters and internal project Champions will showcase how they have made use of EDISON outputs to date to accelerate and expand data science education in their universities and training centres.*"

Prior to the event, Internal Champions and their Mentors have collaborated in the application of EDISON outcomes (i.e. selected components of the EDISON Data Science Framework) and the preparation of the presentations.

**Table 3: First EDISON Champions Conference**

| | |
|---|---|
| Presenters | • University of Bedfordshire (*Marc Conrad*, *Ingo Frommholz*) |

[1] http://edison-project.eu/content/edison-champions-conference

| | |
|---|---|
| | • University of Perugia (*Gianluca Reali*)<br>• Goethe University Frankfurt (*Kim Hee*)<br>• University of Amsterdam (*Adam Belloum*) |
| Results | • Progress on the development of the CF-DS and DS-BoK was very welcomed and was classified as useful.<br>• Application of CF-DS, DS-BoK w.r.t. to individual Champions' DS programmes was demonstrated by Internal Champions to session participants.<br>• Portrayal of expected outcomes and roadmaps related to collaboration with EDISON.<br>• Presentation and discussion of interim MC-DS version.<br>• The knowledge areas were considered as useful concepts to cluster own programs/courses/modules<br>• The BoK was classified as very engineering-oriented. It was recommended to include math and statistics topics.<br>• The (professional) orientation of the individual universities is heterogeneous, so that the concepts can be adapted more or less only in parts.<br>• The plenary was discussing the current difficulties to derive a single DS profession or curriculum. One possible solution would be to specify core modules to be covered by students in the first semesters and specialization on different disciplines in the upper semesters. Professions like "Data Scientist in Medicine", "Data Scientist in Engineering", "Data Scientist in Statistics" were discussed. |

- 2nd EDISON Champions Conference[2] (March 14-16, 2017, Madrid, Spain), Session 1: "*Data Science course design*", Breakout session on course design

**Table 4: Second EDISON Champions Conference**

| | |
|---|---|
| Presenters | • University of Bedfordshire (*Marc Conrad*, *Ingo Frommholz*)<br>• Lucerne School of Information Technology (*Michael Kaufmann*) |
| Results | • How to get students attending your DS courses?<br>• Marketing is important, explaining what DS is and for what reason its valuable to choose this program / course<br>• Renaming courses / combine with familiar terms like "computer" or "information" and set focus<br>• Connection to industry can be attractive for students and useful for industry<br>• Practical parts, like internships or part-time apprenticeship, and flexibility in curriculum would be useful<br>• Case studies and working groups, independent projects to push creativity<br>• Alignment with EDSF is very well received, recognized as a benchmark, to prove to be aligned with DS state of the art<br>• EDSF provides guidance to (re-)create / design DS courses, Learning Outcomes, BoK worked very well |

[2] http://edison-project.eu/content/second-edison-champions-conference

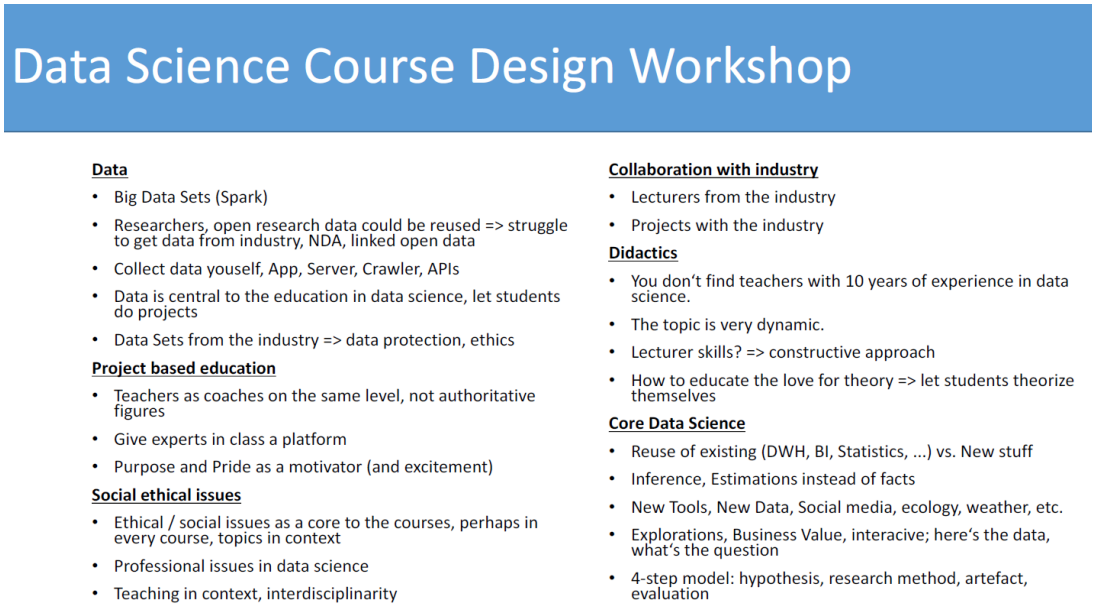



**Figure 1: Second Champions Conference - Course Design Workshop Results**

The results of Session 4: "*Professional pathways: certification and accreditation"* are documented in Chapter 3 Accreditation Use Case (C).

- 3rd EDISON Champions Conference[3] (June, 19-20, 2017, Warsaw, Poland), Session 3 – "*Data Science course design*"

**Table 5: Third Champions Conference**

| | |
|---|---|
| Presenters | • University of Amsterdam (*Adam Belloum*)<br>• University of Stavanger (*Tomasz Wiktorski*) |
| Results | • Data Science can be seen as similar to Medicine education, each physician is educated as a general practitioner and can specialize later.<br>• Knowledge artefacts in Data Science become quickly outdated, it's important to separate competences from knowledge.<br>• It is difficult to educate a Data Scientist without introducing domain knowledge, use practical problems relevant to the group, e.g. Twitter data for young students, 17 sustainable development goals or specific industrial problems (if possible), involve students in grants (research-based education)<br>• Important to differentiate between training and education (professional vs well rounded person w/ more general competences)<br>• Most of DS programs are on MSc level, students come from different BSc and different universities. Often half of the program is used for reeducation what is not efficient. Require basic skills (stats, programming), reeducating shouldn't be more than 10%. |

Building on the initial implementation by Champions and the discussions at the 1st EDSION Champions Conference, the specification of the EDISON Piloting Use Case A scenarios were further refined w.r.t. the systematic application and evaluation of EDISON Data Science Framework components.
To showcase the approach and to prove the usability and manageability, complete mappings to knowledge areas (DS-BoK) and competence groups (CF-DS) were provided for exemplary Internal Champion's program and comprised modules/courses (cf. section 2.3). The initial discussion of the MC-DS took place at the conference.

[3] http://edison-project.eu/third-edison-champions-conference-warsaw-poland

The 2nd Champions Conference discussed the richness of topics and domains where DS plays a role already today and even more in the future. Too much diversification of study offerings was considered uneconomical and, possibly, dissuasive or confusing for students. Focusing study programs on specific priorities could be a solution. Likewise, cooperation with the industry was appraised to tailor DS offerings to the actual need of the market and employers.
Furthermore, our Champions reported back the first positive results of using the EDSF components and promoting the individual DS courses at reviews by proving the alignment with the EDISON Data Science Framework. The design and structuring of new DS programs can be aligned along the Competence Groups and Body of Knowledge components smoothly. Guidance and usability were evaluated as very helpful.

The 3rd Champions Conference presented the new EDSF release. Critical points from the first and second conference were taken up. This has been rated as a very good implementation.

## 2.3 EDSF Application

The engagement with practitioners from academic and training institutions has one decisive advantage: The output of a formal development of a theoretical framework can be validated directly against the real world usage and usability. The only limitation is that testing and implementation in practice do not happen overnight, especially in academic and bureaucratic environments. That is why we can only demonstrate first pilots and document the future plans of our internal Champions.

Collecting the outcomes of the Champions reflection on the EDISON EDSF components, a so-called Illustrated Walkthrough document and a corresponding excel spreadsheet were developed and successfully applied. The detailed results are listed in **Appendix C**. EDSF Application Walkthrough Details.

The input was collected from University of Amsterdam (UvA), Lodz University of Technology (TUL), University of Stavanger (UiS), University of Perugia (UoP), Goethe University Frankfurt (GOE) and Lucerne School of Information Technology (HSLU) and structured by EDSF components CF-DS, DS-BoK, MC-DS and DSPP.

After given a short summary of the existing or planned Data Science programs and courses at the corresponding University, every Champion provided an initial EDSF Reflection.
Afterwards they provided a step-by-step EDSF application:

- CF-DS: Mapping of competence groups competences and learning outcomes
- BoK-DS: Specification of matching knowledge areas and topics
- DSPP: Assignment of professional profiles targeted by program
- MC-DS: Curricula Design and/or Evaluation, Gap analysis

All provided a short summary and outlook section at the end, describing their plans for addressing gaps in the existing offerings or for designing new program(s) complying with EDISON outcomes.

### 2.3.1 Initial Reflection

The initial reflection provides a first impression on assessing scientific and technical soundness, the scope and relevance, the readability, appearance and structure, the usability. In addition, the Champions were asked for their expectations with respect to applying EDSF.

The feedback was positive and the core statements were as follows:

- EDSF is highly relevant, because the need for data science education is growing, the demand in the industry is raising and it is not matched by educational offerings
- Identified need for structure and frameworks to give a guidance in DS education and training
- EDSF technically and theoretically very sound, as BoK is aligned with existing books and specific data science competencies, based on the needs of real world entities by analysis of job offerings
- Applying EDSF will help to identify target DS science profiles for which the graduates match in terms of competencies and skills
- Using EDSF will help to identify gaps in the current DS offerings
- EDSF provides a good thinking framework to design new DS programs

- EDSF component documents are well structured and contain a lot of important, useful and well-illustrated information
- The scope of the provided components is rather broad in range
- To provide support on how to design courses for everybody, a well-explained and easy to read guideline has to illustrate the connection between Competences, Learning Outcomes, Knowledge Areas and the Body of Knowledge. Otherwise the effort and time spend will be relatively high to understand and use the documents efficiently

After collecting this feedback, EDISON generated and released a general introduction and guideline for EDSF implementation, available on the EDISON website [43].

The expectations regarding the application of the EDSF were the following:

- Design a new master program in DS, based to the biggest extent possible on existing courses
- Flexible design of program, to allow students with different backgrounds to participate
- EDSF documents should give clear guidelines how to develop new study program or modify existing one for student education at Bachelor and/or Master level in DS related discipline
- EDSF documents should be the source of information about proportions of courses from different fields
- EDSF should serve as a measure of quality management for designing new master programs
- EDSF will serve as international benchmark to aligning individual DS courses
- EDSF will help to identify and close gaps and thus improve the quality of new/existing DS offerings

### 2.3.2 Execution of Application

Following the four components CF-DS, BoK-DS, DSPP and MC-DS, the participating Champions provided valuable feedback regarding specific DS related courses / programs.

University of Amsterdam (UvA) provided a straightforward example analyzing an existing Master program on **Big Data Engineering** with 5 core courses.

University of Stavanger (UiS) mapped the available information of a newly designed **Data Science** Master program with 3 courses.

Goethe University Frankfurt (GOE) provided mappings regarding two Data Science courses **Big Data Technology** and **Introduction to Data Science**. Both will begin in autumn 2017.

University of Bedfortshire (BEDS) mapped their redesigned courses **Information and Data Systems** and **Computing and Data Science** successfully with EDSF.

Lodz University of Technology (TUL) provided detailed feedback about the existing Bachelor program **Biomedical Engineering** and 4 corresponding courses.

Lucerne School of Information Technology (HSLU) generated a detailed mapping regarding the Master / Major program **Data Engineering and Data Science** with 8 core courses.

Persontyle provided feedback regarding the professional profiles and will implement the EDSF knowledge into the restructuring of Data Science related continuing professional education and training offerings of their School of Data Science.

The whole bunch of mappings is provided in the individual sections in **Appendix C**. EDSF Application Walkthrough Details.

**Feedback on CF-DS**

- Mapping straight forward and logical

- Indicates strong parts and more weak parts of individual DS program
- Course specific LOs are often very specific and directed towards specific elements in BoK relevant for the course, while program LOs are general -> mapping could be difficult
- Course LOs often don't follow Bloom's taxonomy what creates an additional challenge in matching
- Indicates balance of theoretical parts and practical parts of individual DS offerings

**Feedback on BoK-DS**

- Aligning DS courses with the EDSF in terms of competences and learning objectives required some expertise in the field
- Different terminologies and different level of granularity in course descriptions could be a challenge
- Is perceived as a guideline

**Feedback on MC-DS**

- A general challenge is that depending on the department that drives the program there is a push to imbalance of the program towards the competence group closest to departments core competences. This is both due to academic and economic reasons.

**Feedback on DSPP**

- Data science is an emerging field subject to frequent it is thus necessary to revise and adapt Data science curricula to keep them relevant to the DS job market
- EDISION methodology is taking into account DS professional profiles that are defined based on the ESCO occupation family profiles, so the EDISON Data Science Framework can be easily adapted and thus form a good reference for institutions offering DS curricula and trainings to monitor the data science evolution and identify gaps in their Data Science offerings

### 2.3.3 Outlook

UvA stated that the will add a new course addressing Data Management competences to the UVA-VU DS engineering track. UvA staff members in coordination with the Research Data Alliance are defining the content of the course which will be available as a series of trainings on *Research Data Management Literacy* and will be host at some of the RDA plenary meetings and thus available to All EDISON champion universities. The formal adoption of the topic Data Management in the UVA-VU DS engineering track is expected for the next academic years.

UiS recognized the need to update Learning Objectives of individual courses, so that it will be possible to better evaluate them in the context of the whole program.

TUL stated that it would be necessary to introduce large changes in the program of Biomedical Engineering, so that it matches all requirements presented in EDISON documents. All learning outcomes and professional profiles must be rearranged according to EDISON guidelines. More courses related to data analysis, data visualization, data transmission and security must be introduced. Necessary is to include Database course in the program.

HSLU is thinking about including more competences / components of the BoK into their curriculum. Since the program is a specialization of existing BSC IT and business IT, it cannot cover all the topics. However, they are evaluating to implement a Master's program for Applied Data Science and they will use the EDISON guidelines for the curriculum development.

The program has structured the initial concept starting with EDISON professional profiles mapped to competency groups, and have identified possible courses using the EDISON DS-BoK Knowledge Units. However, the program yet to implement the individual learning outcomes knowledge levels. The framework helped them greatly to structure the formidable task of designing a new program from the scratch.

Aligning it with an existing EU framework gives much more credibility to the concept, it simplifies the task, because there is a catalogue of elements to choose from, which ensures kind of a quality standard.

## 2.4 Implementation of Online Training Courses

The EDSION Online Educational Environment (Task 3.4) is described in detail in Deliverable D3.4.
The portal is available at *https://datasciencepro.eu.*

FTK provided, together with Globit, the eLearning and Course Management System for implementation through ENG and EGI.

EDISON started to promote the EDISON Platform and Marketplace to the community at the Champions Conference in Madrid and asked the participants and online spectators to provide input to create, upload and present their own Online Courses.

The Online Educational Environment is therefore offering the operationalization of the EDISON Competence Framework. Following the Data Science Model Curriculum [6] the whole Competence Framework [4] with corresponding Knowledge Areas and Learning Outcomes [5] was implemented and supplemented by the family of Professional Profiles [7].

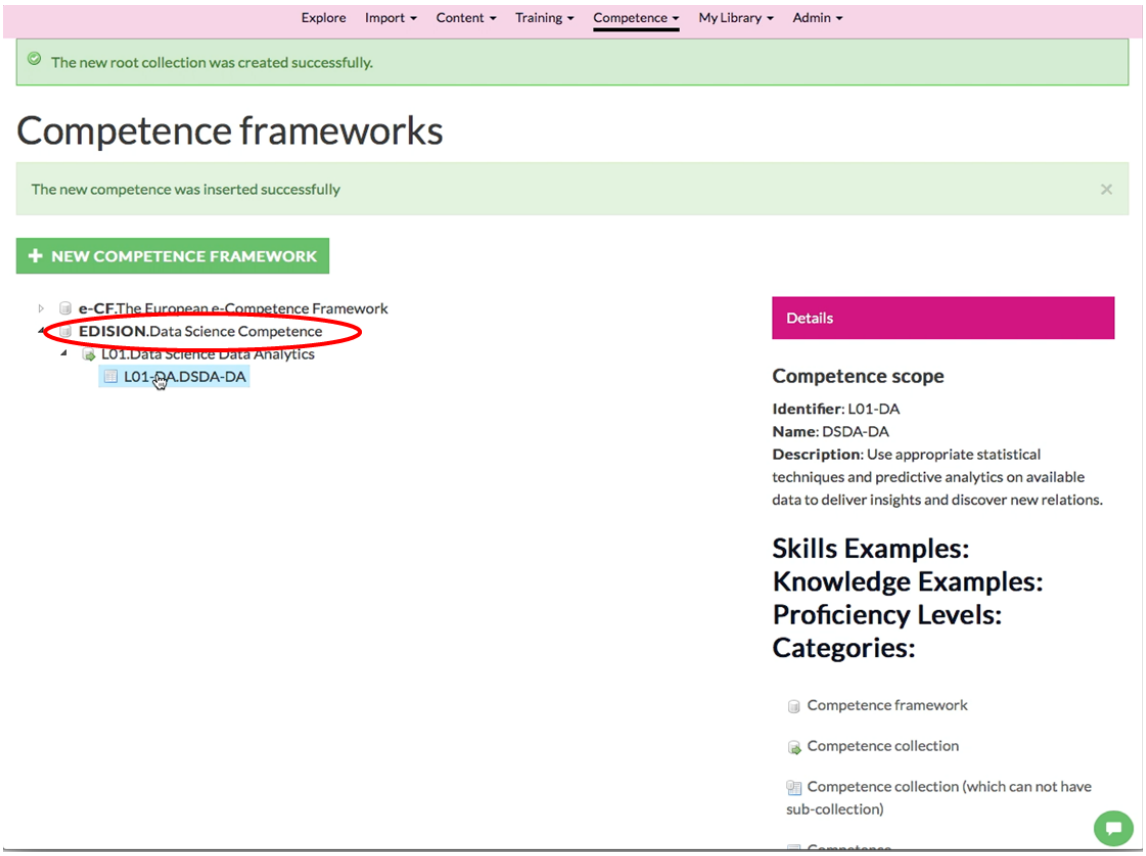



**Figure 2: The online representation of the EDISON Competence Framework**

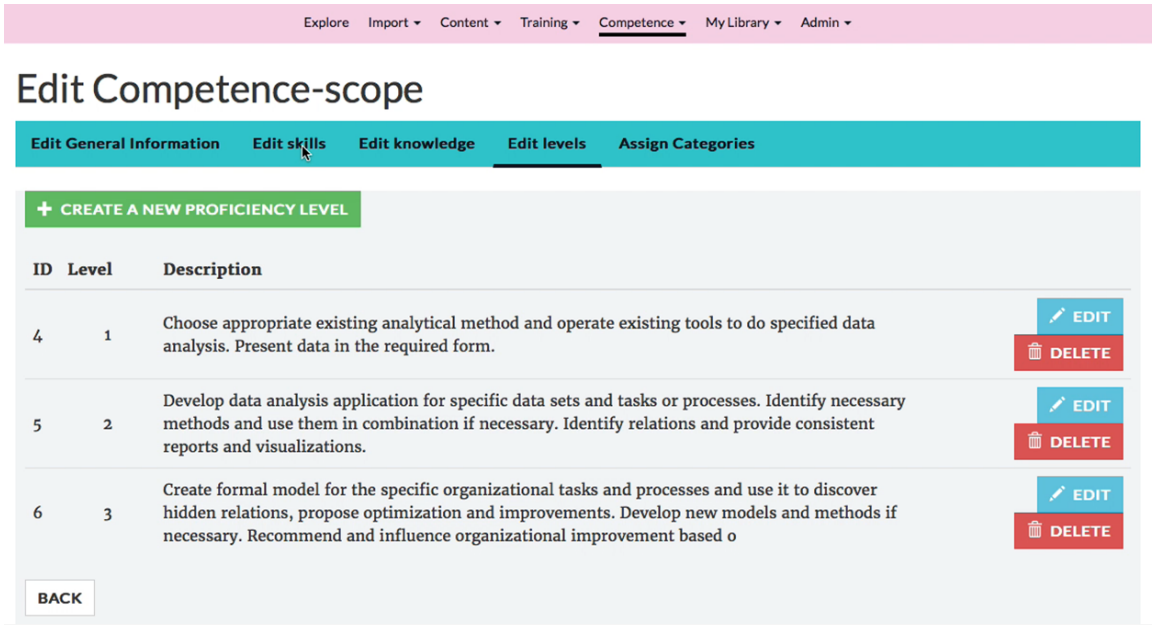



**Figure 3: Create or choose proficiency level**

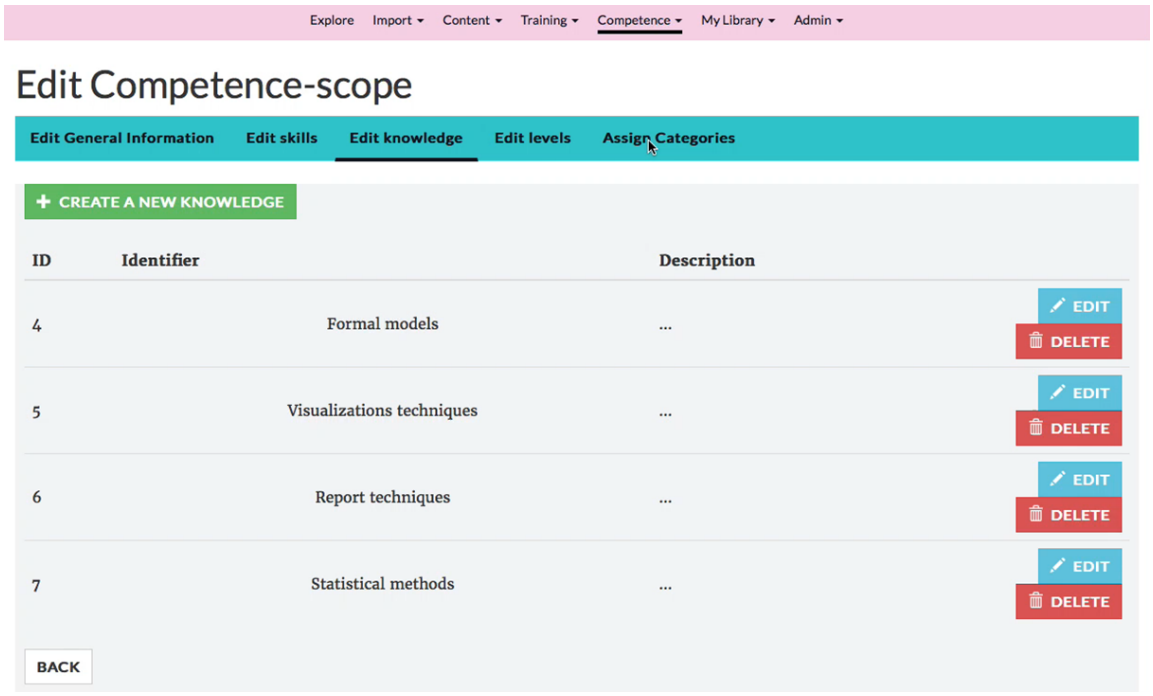



**Figure 4: Example for available knowledge areas**

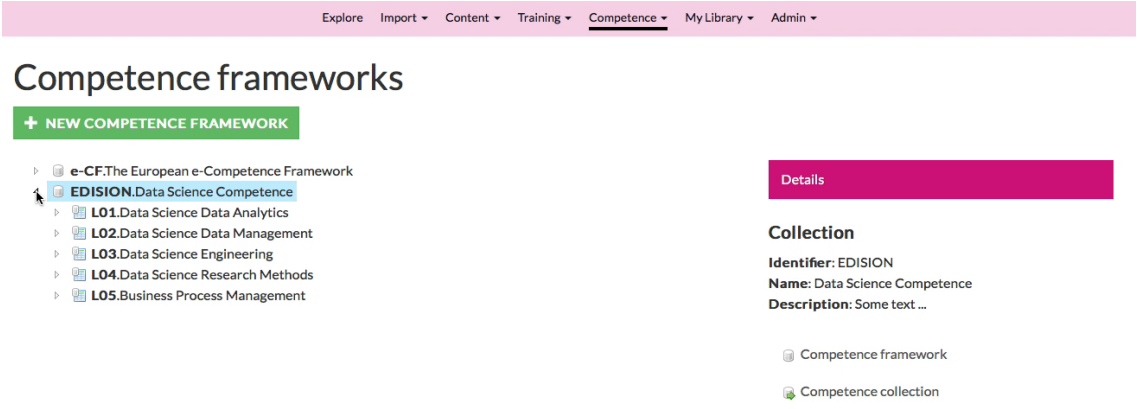



**Figure 5: Example for available Learning Objectives**

Following a step-by-step approach, the first Online Courses were uploaded and validated the system in parallel. Some are provided with an academic / higher education focus and some are provided as Life-Long Learning training course.

**a)** Academia: **High Performance Computing and Big Data** by Adam S.Z. Belloum, University of Amsterdam, The Netherlands

The objective of these courses is to enhance the accessibility of the HPC (High Performance Computing) facilities. The course is composed of a set of lectures covering various topics related to HPC and Big Data and a number of hands on labs and workshops to develop practical skills. The course has been designed to instruct students and researchers who are still unfamiliar with the use of the available HPC facilities or a number of its new components.

**Presentation in the Online Educational Environment:**

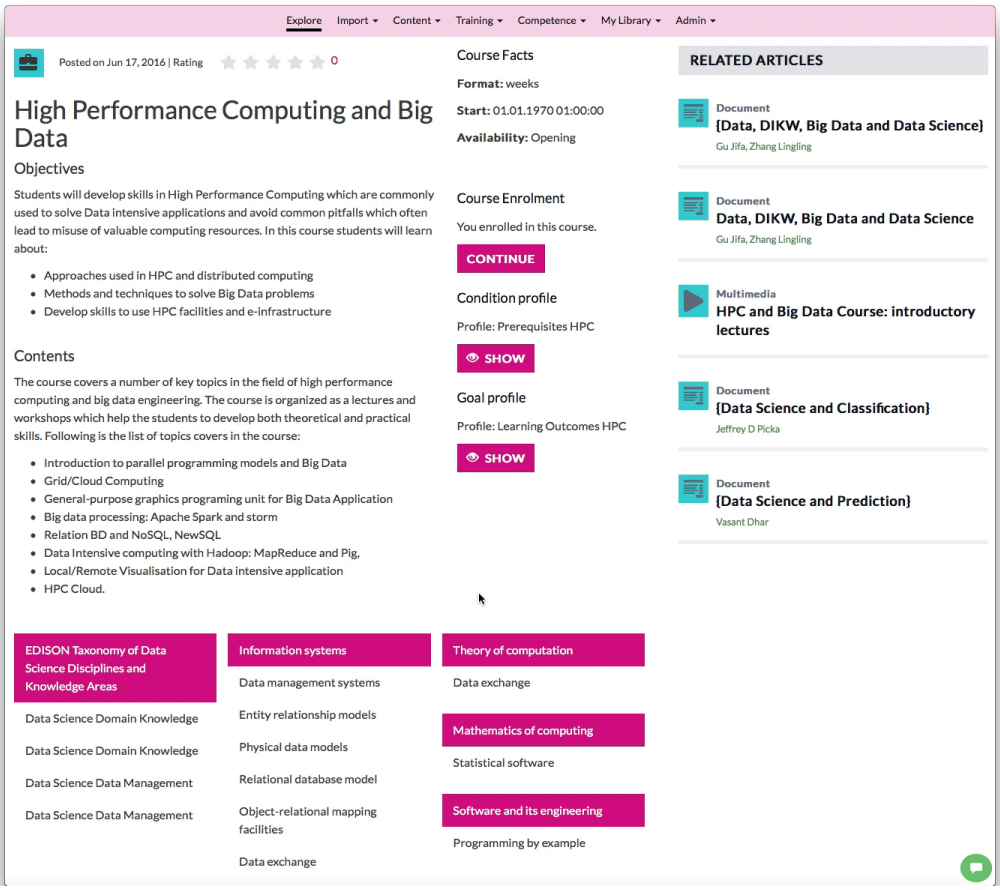



**Figure 6: High Performance Computing and Big Data online course**

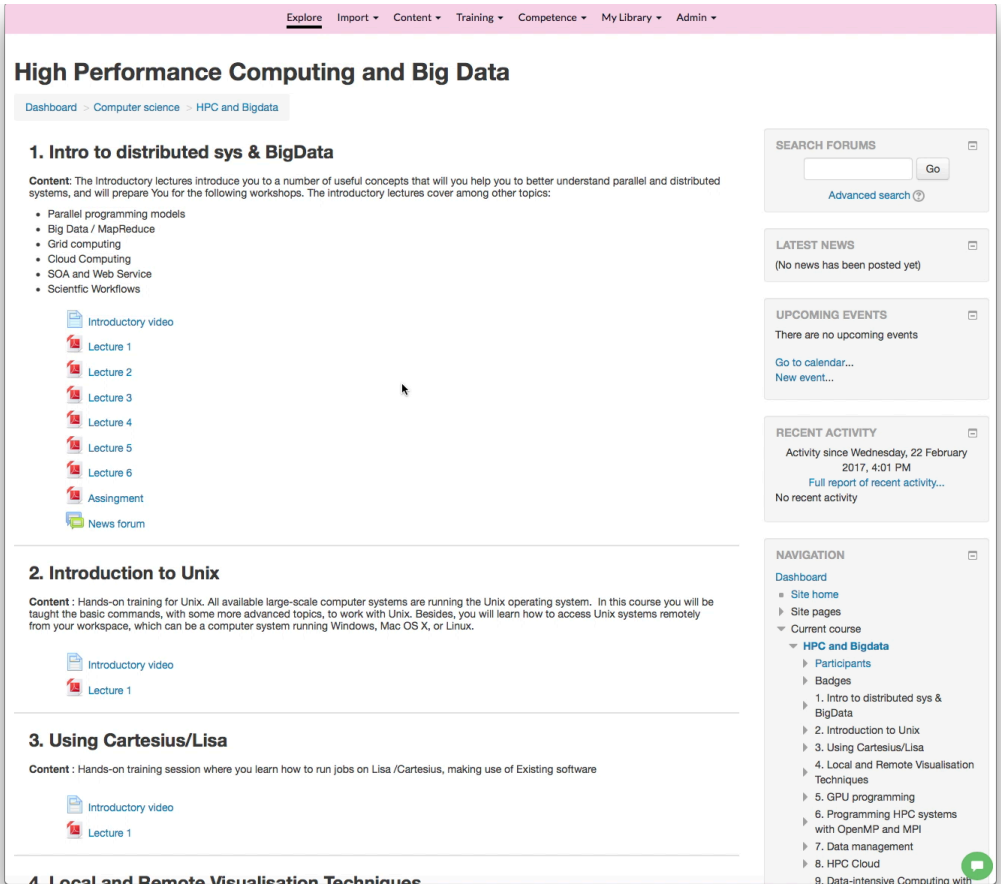



**Figure 7: High Performance Computing and Big Data inside (I)**

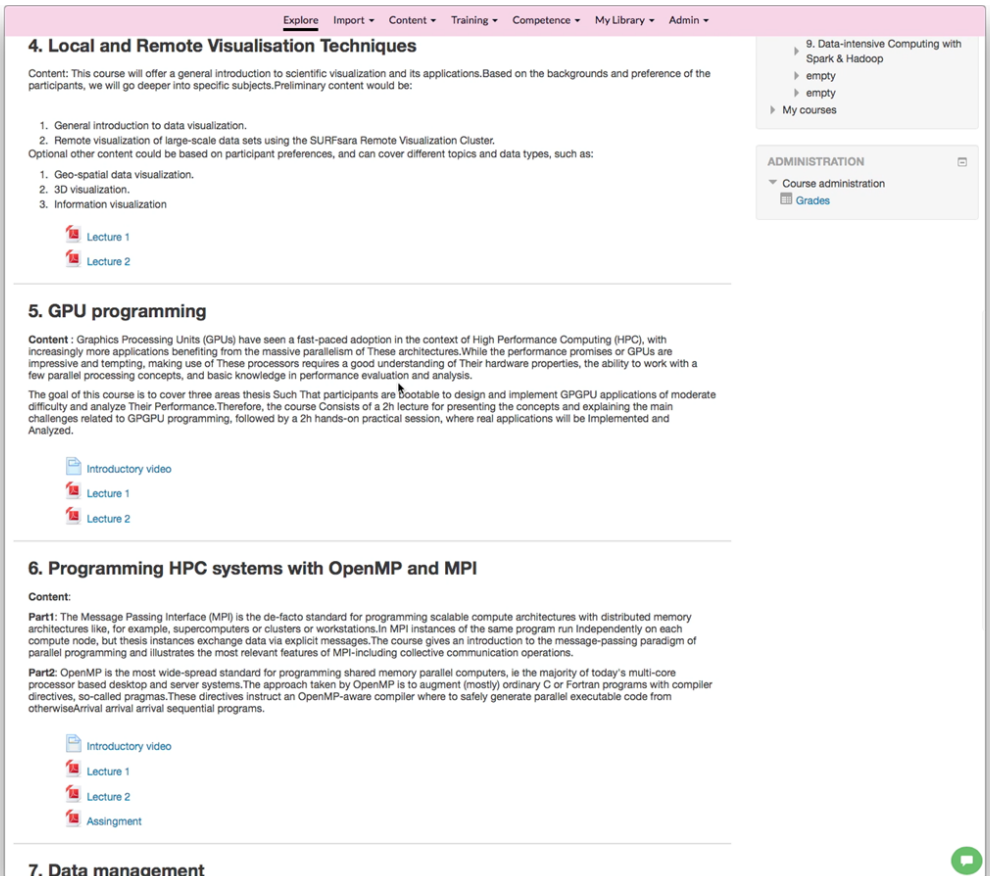



**Figure 8: High Performance Computing and Big Data inside (II)**

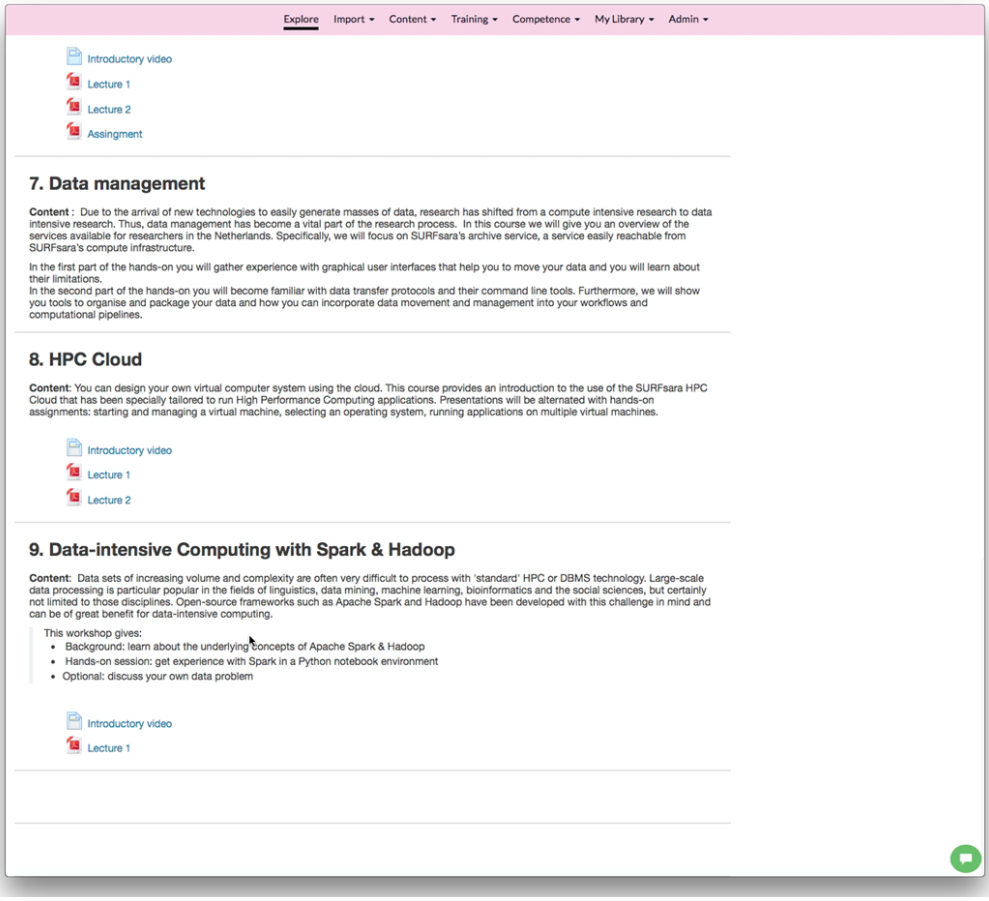



**Figure 9: High Performance Computing and Big Data inside (III)**

**b)** Life-long Learning: **Introductory Course to Data Science: Director Level - Learn and Empower (LARK)** by The Jožef Stefan Institute, Slovenia for EDSA (European Data Science Academy)

The course emphasises a hands-on approach to learning the theoretical knowledge and technical skills to apply the powerful insights of data science, offering a number of exercises that will let the students try out many of the techniques and concepts covered in the taught material. The course is providing the foundations for a career in Data Science.

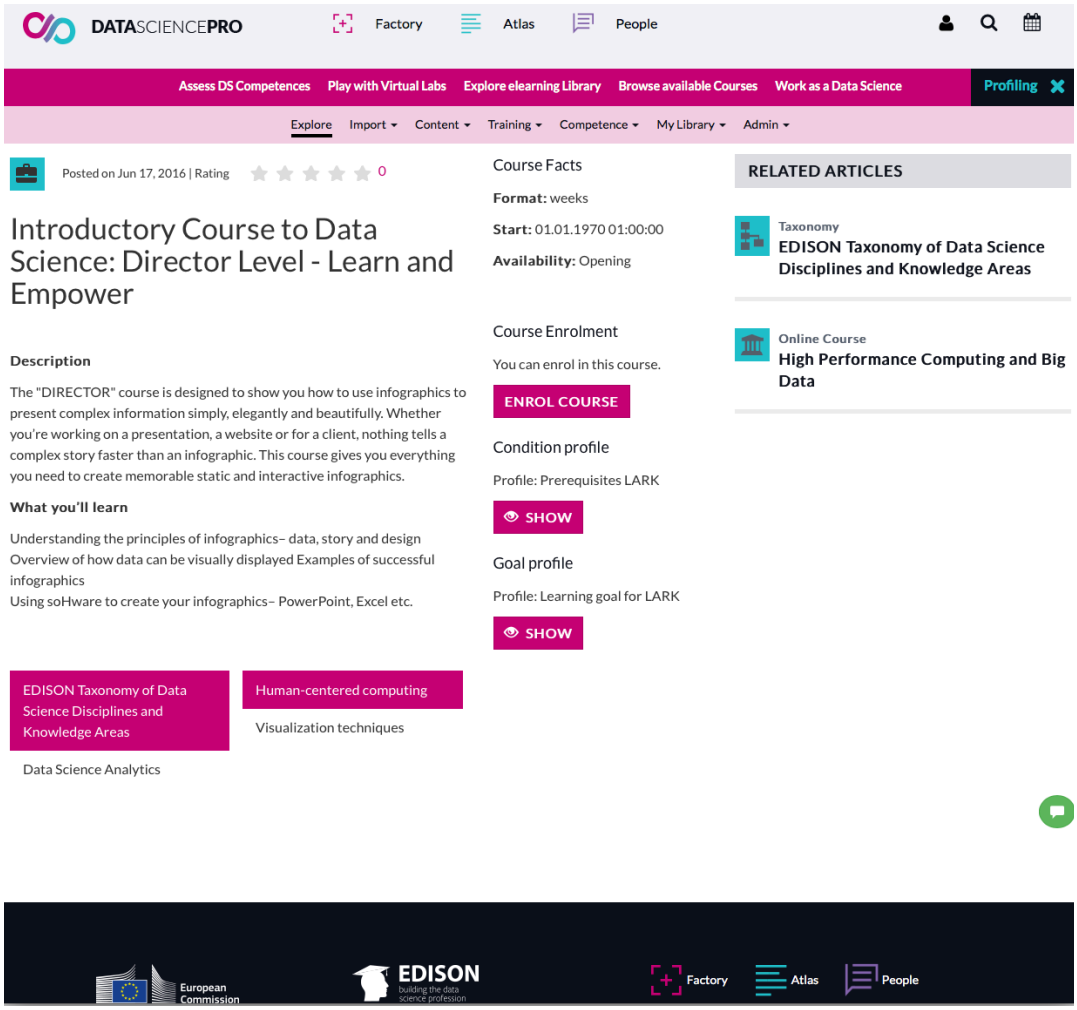



**Figure 10: Introduction to Data Science online course**

**c)** Life-long Learning: **Big Data Analytics** by The Jožef Stefan Institute, Slovenia for EDSA (European Data Science Academy)

This module provides an overview of approaches facilitating data analytics on huge datasets. Different strategies are presented including sampling to make classical analytics tools amenable for big datasets, analytics tools that can be applied in the batch or the speed layer of a lambda architecture, stream analytics, and commercial attempts to make big data manageable in massively distributed or in-memory databases. Learners will be able to realistically assess the application of big data analytics technologies for different usage scenarios and start with their own experiments.

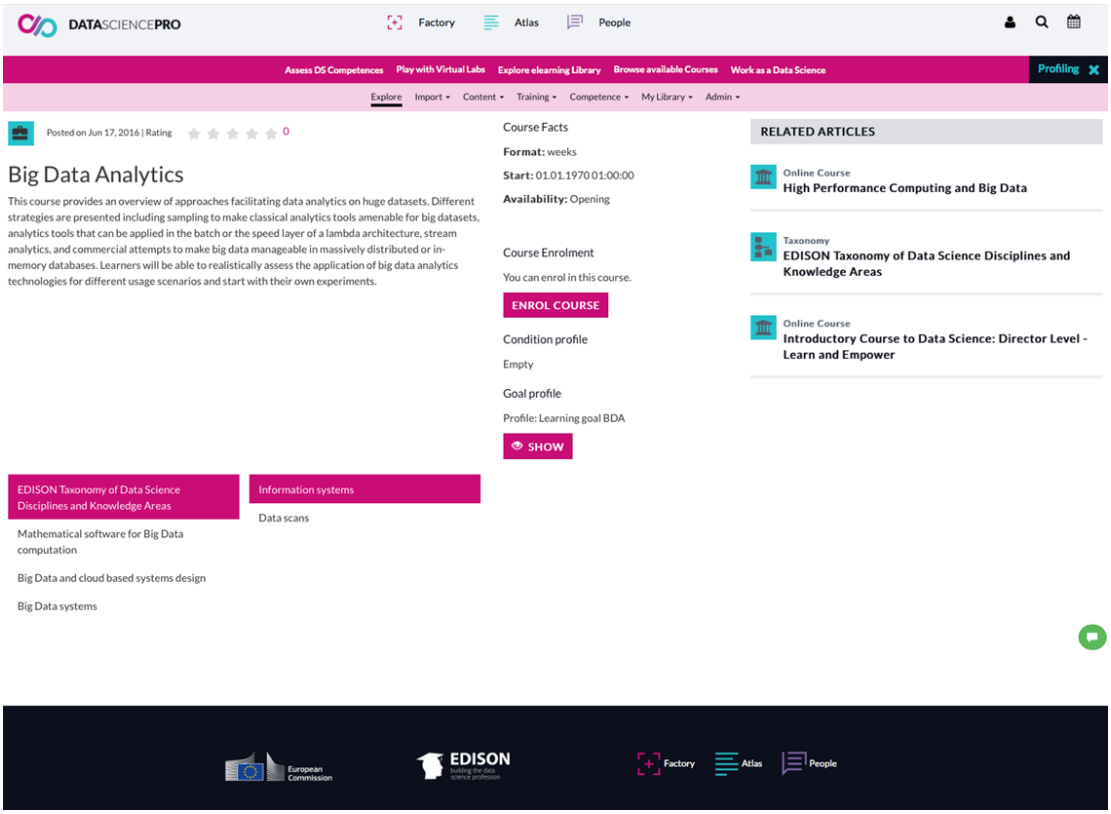



**Figure 11: Big Data Analytics online course**

## 2.5 Further Engagement and Plans

### 2.5.1 Champions Conference

The EDSION Champions Conference was recognized as a place to connect with the Data Science community, to exchange about challenges, changes and progress in the domain and related domains and to get consultancy in using the EDISON Data Science Framework as structured guideline to align current or new Data Science education and training offerings und professional profiles.

The implication of the EDISON project will get visible in the next years. University of Bedfortshire, Lucerne School of Information Technology, University of Perugia, University of Amsterdam and also The School of Data Science and the Engineering Training Center and many others are currently developing new Data Science offerings or restructuring running Data Science education and training offerings and want to align them with the EDSION Framework.

Furthermore, the discussions beside the formal sessions and in smaller groups at the Champions Conference were very fruitful. The ideas on accreditation, certification, batches, interdisciplinary courses and co-operations only just started to grow and many of the participants would like to continue the exchange with likeminded colleagues and even with new colleagues from different countries, to maintain and expand the contacts with the Data Science community.

To address this positive feedback and the further developments and to get a continuity of a hitherto valuable event, the interested EDSION consortium members, with the leadership by FTK and initiative group involving University of Southampton and University of Amsterdam, decided to continue the Champions Conference as an independent event beyond the runtime of the EDISON project.

The Champions Conference will keep promoting the EDISON legacy, will provide a platform for the Data Science community to present their work, discuss progress and best practices and will give advice to a broad range of Data Science related disciplines.

The next Champions Conference is planned to take place in Edinburgh, Scotland in May 2018. The Conference will be organized by PCO FTK e. V. with Prof. Matthias Hemmje and Steve Brewer, University of Southampton, as Organizational Chairs. Co-Organizer will be the Data Lab in Edinburgh, together with the University of Edinburgh. University of Amsterdam will act Program Committee co-chair.
The Conference will be sponsored the Data Lab Edinburgh[4] (Joshua Ryan-Saha), The Alan Turing Institute[5] (Gian Marco Campagnolo) and expected to receive sponsorship from IBM, Big Data & Analytics Hub[6] and potentially Oracle.

### 2.5.2 Online Courses

In cooperation with the Champion Universities, Persontyle, Engineering Training Center, Jožef Stefan Institute and EDISON partner EGI and FTK the portfolio of online courses will be extended. Agreements have already been made to expand the range in the coming months.

Examples for the next courses to be available online are the following:

d) **Information Visualization online** by Matthias Hemmje, FernUniversität Hagen, Germany
e) **Information and Knowledge Management online** by Matthias Hemmje, FernUniversität Hagen, Germany
f) **Data Management** by Adam S.Z. Belloum, University of Amsterdam, The Netherlands
g) **Big Data Analyzing in EGI Fed-Cloud using Hadoop Infrastructure (working title)** by FTK/EGI
h) **Fundamentals of Machine Learning** by Persontyle, UK

The Champions Conference will be used to promote the EDISON portal to interested stakeholders and provide the possibility for guided instructions how to create online courses independently.

[4] http://www.thedatalab.com/
[5] https://www.turing.ac.uk/
[6] http://www.ibmbigdatahub.com/

# 3 Accreditation Use Case (C)

As a follow up of the preliminary work on accreditation described in Deliverable D3.1, the EDISON consortium took some steps toward defining the expected impact on Accreditation of the Data Science programs. The preliminary study showed large discrepancy in the accreditation procedures of new programs across different EU countries. When talking to various accreditation agencies it become clear that the accreditation landscape across EU countries and accreditation agency is very diverse. There are several approaches to accreditation in different EU countries: program accreditation, institution accreditation of self-accreditation.

**Accreditation Session during EDISON Champions workshop**

One of the mechanisms followed in EDISON to structure and bring the various stockholders around the table to discuss the accreditation topic was to organize sessions during the EDISON Champions Conferences. Representatives of both EU universities and some accreditation agencies attended the designated accreditation sessions.

- The **first session on accreditation and certification was held in Madrid during the 2nd EDISON champion conference** [33] [34], where we invited a representative from the Dutch Accreditation Agency to inform the participants about the common procedures of program accreditation in The Netherlands and Belgium [35].
- The **second session was held during the 3rd EDISON champion conference in Warsaw** [36], where we invited the chair of working group responsible for accreditation and certification processes in Conference of Rectors of Academic Schools in Poland which is the body that coordinates cooperation of member institutions and represents common interests of the schools as well as undertakes activities leading to the creation of effective and integrated system of national education. The focused discussions during both events created a discussion platform which helped to clarify various existing accreditation practices to the participants and helped them to reflect on their respective accreditation cases and possibility to further investigate the appropriate approach to be accepted by their university and country.

Next to the activities on Accreditation organized in the context of the last two EDISON champions conference, EDISON partners engaged with National accreditations namely the Italian and Spanish accreditation agencies:

**ANVUR** (Agenzia Nazionale per la Valutazione delle Università e della Ricerca) [37] the Italian agency in charge of the accreditation of new course and degrees based on ex-ante and ex-post evaluation of the overall quality of the Academic offering.

As part of this process, each Institution and each degree program (“corso di laurea”) need both an initial ex-ante accreditation and a periodic ex-post accreditation (each five years for the institution and each three for each program). Such evaluation is done by the National Evaluation team based on the information collected by a specific quality team in each University (that acts as collector and unique contact point for the National Evaluation team).

The normalisation with a common set of indicators of the accreditation allows comparison and trends analysis among the Universities and within each of them. The full guideline is available on the ANVUR website. However, some indicators may be related to the EDISON activities, because they indicate the needs for Scholars and Administrative personnel in the Universities, for which EDISON may provide support.

In particular the indicator R3 is focusing on the Degree Program. It is composed of four set of information that needs to be provided either at the initial and the periodic evaluation stage:

- R3.A – the DP defines the cultural and professional profiles of the person intends to educate to and propose coherently training activities.
- R3.B – the DP promote a student-centered didactics and promote the adoption of updated methodologies and verify correctly the acquired competences
- R3.C – the DP has an adequate team of scholars and technicians offer service accessible to students and exploit locations coherent to the didactic needs.
- R3.D – the DP is able to identify critical aspects and ways of improvement of internal educative organisation and it is able to implement actions consequently.

**ANECA (**Agencia Nacional de Evaluación de la Calidad y Acreditación ) [38] is the Spanish accreditation agency. ANECA sets a number of protocols and rules for program accreditation. For a program to be accredited by ANECA it must have be relevant to teaching or research experience, adequate to the academic field of reference, and in accordance with similar existing international studies.

- The general objectives and competences of the degree have to be justified, according to the academic field. Competencies have to be assessable.
- The ways of access to the studies have to be explicit, and teaching content has to be planned, and the adequate teaching staff and human resources have to be allocated.
- The proposal has to include foreseen results, and there has to be a quality guarantee system, that is, procedures associated to quality guarantee; and there have to be some formal mechanisms for approval, control, regular reviewing and improvement.

Next to the programs accredited by ANECA, there are the universities' own program. In general, the universities' own degrees have a lower number of credits, and are shorter studies. Normally universities start by having their own degrees (typically master's degrees) and if they want to have an "accredited" degree the process takes some 2 to 4 years.

The EDISON investigation about the data science program accredited in Spain showed that most of the existing DS programs (expect one) are owned by universities and thus are not yet accredited by ANECA

**Table 6: European Approach for Quality Assurance of Joint Programmes**

| **Name of studies [university]** | **Accredited by ANECA** |
|---|---|
| Máster en Business Data Science[7] [Universitat Jaume I] | No |
| Experto en Data Science[8] [U-TAD] | No |
| Máster en Data Science[9] [Universidad Rey Juan Carlos]<br>Máster en Big Data y Business Intelligence[10] [Universidad Rey Juan Carlos.] | No<br>No |
| Máster en Big Data y Data Science: Ciencia e Ingeniería de Datos[11] [Universidad Autónoma de Madrid] | No |
| Máster en Data Science[12] [Universidad de Alcalá] | No |
| Máster en Big Data Science[13] [Universidad de Valladolid] | No |
| Máster Universitario en Big Data y Data Science[14] [Universidad Internacional de Valencia] | In progress |
| Ciencia de los datos (Data Science)[15] [Universitat Oberta de Catalunya] | Not specified |

Overview of existing accreditation system in Europe

Because in EDISON we are targeting Data Science educational programs, which are inherently multi-disciplinary and requires a lot of human and infrastructure resources to set up, participants of the first EDISON Champions Conference, in particular session on Professional pathways: certification and accreditation, were interested in the creation of joint programmes across EU countries. This was the topic of the representative from the Dutch Accreditation Agency. The talk pointed out that despites the different of national Qualification Assurance

[7] http://www.fue.uji.es/masterbusinessdata/all
[8] http://www.u-tad.com/ld/17-18/postgrado-experto-data-science/?gclid=Cj0KEQjwkZfLBRCzg-69tJy84N8BEiQAffAwqud_ydz9bCiShC3IbAfcvbleS50wFFDqHD03YaBLAMkaAurV8P8HAQ
[9] http://www.masterdatascience.es/
https://www.urjc.es/estudios/titulos-propios/1332-data-science
[10] http://www.eadic.com/cursos/agua-energia-medioambiente/master-en-big-data-y-business-intelligence/?gclid=Cj0KEQjwkZfLBRCzg-69tJy84N8BEiQAffAwqpeNXuBgf5g_NiPVbx0R5vCrU6TRRzfiIi9_N0C7BG0aAjQV8P8HAQ
[11] https://www.uam.es/ss/Satellite?c=UAM_EstudPropio_FA&cid=1242692602422&language=es&pagename=UniversidadAutonomaMadrid%2FUAM_EstudPropio_FA%2FUAM_estudioPropio&pid=1242654675830&title=M%C3%A1ster+en+Big+Data+y+Data+Science%3A+Ciencia+e+Ingenier%C3%ADa+de+Datos
[12] https://www.uah.es/es/estudios/Master-en-Data-Science/
[13] http://www.tel.uva.es/docencia/planes/masterbds.htm
[14] http://www.viu.es/master-universitario-big-data-data-science/
[15] http://estudios.uoc.edu/es/masters-posgrados-especializaciones/especializacion/informatica-multimedia-telecomunicacion/ciencia-datos-data-science/presentacion

regimes which have fragmented assessments, multiple procedures and use different frameworks to organize accreditation of program (visits, panels, reports, decisions) there was in 2012 the so-called Bucharest Communiqué which aimed at recognising quality assurance decisions of EQAR-registered agencies on joint and double degree programmes [39]. As a follow up a European Approach for Quality Assurance of Joint Programmes was adopted in May 2015 by EHEA Ministers in Yerevan.

Under the European approach, if some of cooperating High Educational Institutes (HEI) require programme accreditation/evaluation then HEIs should select a QA agency registered in EQAR [40]. Agency will use Standards and Procedure to carry out a single procedure of the entire joint programme; the decision to be recognised in all countries where the programme is offered European Approach may be used by self-accrediting HEIs and countries outside of EHEA [41]. The European approach to joined program defines Standards for 9 different aspects of Educational programs ranging from eligibility of the institutions that could apply for the joined program to the way we define Quality assurance. It also defines a clear procedure for the accreditations composed on 9 steps: (1) Self-Evaluation Report, (2) Review Panel, (3) Site Visit, (4) Review Report, (5) Formal Outcomes and Decision, (6) Appeals, (7) Reporting, (8) Follow-Up, and (9) Periodicity.

European approach to joint program was leading to long discussion during the Champions Conference sessions on accreditation. One of the questions that has been put forward is the role of the EDISON project in the Accreditation of Data Science program. It was clear from the various discussions that EDISON can have a positive impact on the accreditation of new Data Science program, for instance EDSF framework can be used by the various accreditation agencies to identify the expertise that are needed to for the review during the accreditation process. Using EDSF, the accreditation agencies can form review panel with right expertise. In this sense, the few contacts the EDISON consortium has managed to establish with accreditation agencies of the countries participating in the consortium has created an awareness on both sides (HEI and Accreditation agency) about the complexity of establishing and accrediting data science program.

The EDISON champion universities might end up trying to establish such a joint European Data Science program. This might take some time to materialize, it was a clear take away message of the Madrid Session on accreditation: joint programs are possible in theory but are difficult to put in practice. A few experiences of recently joint program accredited by Dutch accreditation agency in 2016, have been presented and discussed. An informal initiative has been realised during the course of the project where the teaching staff from the University of Amsterdam gave a complete course on Data Science Analytics methods at the University of Perugia as a part of in the context of the newly established Data Science master program.

# 4 Summary and Outlook

The work package WP3 - Development and Reference Implementation Strategy has the goal to connect the theoretical results achieved in work package WP2 into concrete actions in the form of use cases for the target communities.

The EDISON Data Science Framework was adopted in support of the piloting use cases by involving Champions from academic and scientific communities as well as training providers and industry.

The Champions participated to assess, align, and adopt EDISON Data Science Framework components and therefore support the design and development of Data Science Education and Training content during EDISON and, especially in the long term, after the runtime of the project.

The Use Cases were performed successfully by the following steps:

- Engagement
- Validation
- Implementation
- Continuation

The engagement events, like the Champions Conference and consequently hands on training and bilateral discussions were very productive. The resonance and the feedback were positive, because the workshops and talks were perceived as informative and future-driven, but especially the encompassing exchange across several Data Science related disciplines was noticed as valuable.

The validation of the EDISON Data Science Framework was very supportive to balance the theoretical ideas and concepts against the usability and practical realization.

The implementation of the EDSF components conducted by the application of the Model Curriculum and the consideration of the framework as a European guideline is well on its way. EDSF was recognized by the target communities and stakeholders to ensure usability and quality for the design and development of new education and training offerings in the important branch of Data Science.

Furthermore, the project legacy will be continued and this is the most important part after the initial positive feedback and proved usability. The sustainability will be supported by establishing regular Data Science Conference primarily focused on the European academic and research community and industry.

# 5 References


[1] EDISON Description of Action – ANNEX 1 (parts A, B), Coordination and support action, Number 675419.

[2] EDISON Deliverable D2.2 Existing educational and training resources inventory and analysis, May 2016

[3] EDISON Deliverable D2.3 EDISON Data Science Framework (final version), June 2017

[4] Data Science Competence Framework [online] http://edison-project.eu/data-science-competence-framework-cf-ds

[5] Data Science Body of Knowledge [online] http://edison-project.eu/data-science-body-knowledge-ds-bok

[6] Data Science Model Curriculum [online] http://edison-project.eu/data-science-model-curriculum-mc-ds

[7] Data Science Professional Profiles [online] http://edison-project.eu/data-science-professional-profiles

[8] EDISON Deliverable D3.1 Model Curricula definition and Report on the Use Cases support design, August 2016

[9] EDISON Deliverable D3.2 EDISON Online Education Environment, September 2016

[10] European Skills, Competences, Qualifications and Occupations (ESCO) [online] https://ec.europa.eu/esco/portal/home

[11] University of Bedfordshire - Data Science - BSc (Hons) programme [online] http://www.beds.ac.uk/howtoapply/courses/undergraduate/next-year/data-science

[12] University of Perugia – Master Certificate in Data Science [online] http://masterds.unipg.it/

[13] Goethe University Frankfurt – Faculty of Compuzter Science and Mathematics [online] http://www.goethe-university-frankfurt.de/58074304/Faculty-of-Computer-Science-and-Mathematics

[14] Goethe University Frankfurt - Frankfrut Big Data Lab [online] http://www.bigdata.uni-frankfurt.de/teaching/

[15] Amsterdam Data Science [online] http://amsterdamdatascience.nl/

[16] Amsterdam Data Science – Education [online] http://amsterdamdatascience.nl/education/

[17] Vrije Universiteit Amsterdam – Computer Science Specializations [online] http://vu.nl/nl/opleidingen/masteropleidingen/opleidingenoverzicht/c-d/computer-science/specializations/index.aspx

[18] University of Amsterdam/SURFsara - HPC and BigData course [online] http://hpc.uva.nl/

[19] University of Stavanger – Course Data-intensive Systems [online] http://www.uis.no/studies/study-courses/?categoryID=10648&parentcat=9835&code=DAT500_1&name=Data-Intensive+systems

[20] University of Lodz – Specialization Biomedical Engineering [online] http://www.programy.p.lodz.pl/kierunekSiatka.jsp?l=en&w=Biomedical Engineering&p=4291&stopien=first-cycle programme&tryb=full-time

[21] Engineering Training Center "Enrico Della Valle" [online] http://www.eng.it/lavora-con-noi/it-academy.dot

[22] Persontyle Data Science Centre of Excellence [online] http://www.persontyle.com/school/

[23] Persontyle Data Science Centre of Excellence – Learning Programs [online] http://www.persontyle.com/courses/

[24] Lucerne University of Applied Sciences and Art – Lucerne School of Information Technology [online] https://www.hslu.ch/en/lucerne-school-of-information-technology/

[25] Lucerne University of Applied Sciences and Art [online] https://www.hslu.ch/en/

[26] Lucerne University of Applied Sciences and Art – Degree Programmes [online] https://www.hslu.ch/en/lucerne-school-of-information-technology/degree-programs/bachelor/majors/

[27] Jožef Stefan Institute [online] https://www.ijs.si/ijsw/JSI

[28] EDISON Project: University Programs list [online] http://edison-project.eu/university-programs-list

[29] EDISON Project Documents Library [online] http://edison-project.eu/library

[30] ENQA - European Association for Quality Assurance in Higher Education [online] http://www.enqa.eu/

[31] EQANIE - the European Quality Assurance Network for Informatics Education [online] http://www.eqanie.eu/

[32] Champions Conference I, United Kingdom http://edison-project.eu/content/edison-champions-conference

[33] Champions Conference II, Spain http://edison-project.eu/content/second-edison-champions-conference

[34] Champions Conference II, Spain, Video of the Session 4 – Professional pathways: certification and accreditation: https://arcamm.uc3m.es/arcamm_3/item/show/e4aa254895833d499181d93734591a1b

[35] Champions Conference II, Spain, Slides Lineke van Bruggen (Dutch Accreditation Agency): http://edison-project.eu/sites/edison-project.eu/files/attached_files/node-490/lineke-2017-16-03-madrid-final.pdf

[36] Champions Conference III, Poland http://edison-project.eu/content/third-edison-champions-conference

[37] ANVUR - Agenzia Nazionale per la valutazione delle università e della ricerca, Italien accreditation agency http://www.anvur.org/attachments/article/1039/LG_AVA_2017_05_05.pdf

[38] ANECA - Agencia Nacional de Evaluación de la Calidad y Acreditación, Spanish accreditation agency https://www.mecd.gob.es/prensa-mecd/actualidad/2015/12/20151211-aneca.html, http://www.aneca.es/Programas-de-evaluacion/ACADEMIA

[39] EQAR - European Approach for Quality Assurance of Joint Programmes http://www.eqar.eu/topics/joint-programmes

[40] EQAR Register - European Quality Assurance Register for Higher Education https://www.eqar.eu

[41] EHEA - European Higher Education Area And Bologna Process https://www.ehea.info

[42] Alistair Cockburn - Use case fundamentals [online] http://alistair.cockburn.us/Use+case+fundamentals

[43] EDSION General Introduction to EDSF [online] http://edison-project.eu/sites/edison-project.eu/files/attached_files/node-488/edison-general-introduction-edsf.pdf

# Appendix A. Internal Champions' Terms of Reference

The roles and responsibilities of acting "EDISON Internal Champions" are characterized by the following activities:

- To provide feedback to CF-DS reports (Documents D2.1 due in Month 6 and D2.3 due in Month 22 according to the EDISON workplan);
- To contribute to the preparation of the DS-BoK from Month 6 to Month 22;
- To support the adoption of a MC-DS within the institution (informative part) from Month 12 to Month 24;
- To endorse EDISON legacy by being presenter of "How his/her institution benefitted from EDISON" in at least one EDISON promoted event – an "EDISON-do-it-Better" statement;
- To share learning experience, content and teaching material under a CC-BY-NC-SA license.

From their participation in the project, the "EDISON Internal Champions" will benefit from:

- Improved support in adjusting existing courses content and access to shared material;
- Better justification for the changes they want to do on the educational programme, based on the results of demand analysis;
- Get a closer look and deeper understanding of the labour market place for graduates of Data Science programme and related courses;
- Increased visibility of own training offering and expanded branding opportunities for its Data Science programme and related courses;
- Experts' support in completing the accreditation schema (content and reference numbers from the market analysis);
- Privileged exploitation of the Online Educational Environment (for both promoting Champions courses programs – a sort of "sponsor" link made in evidence – and using it in Labs/Exercises);
- Having costs (Conference Fee – if any, travel and accommodation) covered for promotional activities in which EDISON collaboration and Use Case experience are presented in EDISON selected events (pending prior approval);
- Priority in being invited to EDISON follow-up actions/project/grants.

Both, EDISON and its champions, agree that they will clearly label the source of all content and require users to indicate those sources in any subsequent re-use. Furthermore, they agree to fully acknowledge the other (verbally and by including logos) when citing or promoting jointly developed products under this Cooperation scheme.

Each "EDISON Internal Champion" is mentored by an assigned EDISON partner.

## Appendix B. Memorandum of Cooperation

# MEMORANDUM OF COOPERATION

Between
**"EDISON Mentor"** (A)
And
**"Internal Champion"** (B)



## Introduction

This Memorandum of Cooperation (MoC) provides the basis for the **EDISON Mentor** and **EDISON Champion** (hereinafter 'the Parties') to collaborate within the context of the so-called EDISON project, to develop and share relevant information and other resources needed for the Project, and to share results of such endeavour which contribute to better fulfil their respective institutional mandates.

## The Parties

### EDISON Mentor

...

Internal Champion

...

## Objective of this Memorandum of Cooperation

The objective of this MoC is to formalise a framework for cooperative and collaborative work between the Parties while recognising their respective independent yet complementary missions as summarized under article 2 above.

The **Internal Champion** will act as an "**EDISON Champion"**, its role being characterized by the following activities:

- To provide feedback to CF-DS reports (Documents D.2.1 due in Month 6 and D.2.3 due in Month 22 according to the EDISON workplan);
- To contribute to the preparation of the DS-BoK from Month 6 to Month 22;
- To support the adoption of a MC-DS within the institution (informative part) from Month 12 to Month 24;
- To endorse EDISON legacy by being presenter of "How his/her institution benefitted from EDISON" in at least one EDISON promoted event – an "EDISON-do-it-Better" statement;
- To share learning experience, content and teaching material under a CC-BY-NC-SA license.

From its participation in EDISON works, the **EDISON Champion** will benefit from:

- Improved support in adjusting existing courses content and access to shared material;
- Better justification for the changes they want to do on the educational programme, based on the results of demand analysis;
- Get a closer look and deeper understanding of the labour market place for graduates of Data Science programme and related courses;
- Increased visibility of own training offering and expanded branding opportunities for its Data Science programme and related courses;
- Experts' support in completing the accreditation schema (content and reference numbers from the market analysis);

- Privileged exploitation of the Online Educational Environment (for both promoting Champions courses programs – a sort of “sponsor” link made in evidence – and using it in Labs/Exercises);
- Having costs (Conference Fee – if any, travel and accommodation) covered for promotional activities in which EDISON collaboration and Use Case experience are presented in EDISON selected events (pending prior approval);
- Priority in being invited to EDISON follow-up actions/project/grants.

## Conditions of Use

Both Parties agree that they will clearly label the source of all content and require users to indicate those sources in any subsequent re-use. Each Party agrees to fully acknowledge the other (verbally and by including logos) when citing or promoting jointly developed products under this Cooperation scheme.

## Amendments and Modifications

Either Party may recommend amendments to this Memorandum of Cooperation by notifying the other Party in writing and, with the subsequent agreement of the other Party, then implementing those amendments.

Notwithstanding the foregoing, both Parties undertake to jointly review and revise (as necessary) this Memorandum of Cooperation 2 years after it comes into force.

## Non-exclusive Agreement

This Memorandum of Cooperation is non-exclusive, and in no way restricts each Party from participating in similar activities or arrangements with other public or private initiatives, organizations, or individuals.

1. **Contact Information**

**EDISON Mentor**
**…**

**EDISON Champion**
representative
Position
Address
Telephone
Fax
E-mail

_____ Date:
(EDISON Mentor)

_____ Date:
(EDISON Champion)

## Appendix C. EDSF Application Walkthrough Details

### EDSF Illustrated Walkthrough Reporting Template

UC A/B: ABC
Internal Champion XXX
Mentor: YYY

- Prerequisites
  - Competences and skills framework for Data Science (CF-DS)
    http://edison-project.eu/data-science-competence-framework-cf-ds
  - Data Science Body of Knowledge (DS-BoK)
    http://edison-project.eu/data-science-body-knowledge-ds-bok
  - Model Curriculum in Data Science (MC- DS)
    http://edison-project.eu/data-science-model-curriculum-mc-ds
  - EDISON_DSP-profiles
    http://edison-project.eu/data-science-professional-profiles-definition-dsp
- Summary of Data Science program and courses

Detailed, supplementary data in accompanying spreadsheet. Short summary here...

- Contact information (university/company, unit, contact person with email)
- General information about program
  - Designing a new Data Science program or course or analyzing/improving an existing one?
    - New DS program in the making? What have you already achieved, what are you planning to achieve, what is the current status of organizing / initiating new DS programs / courses?
  - What is the department that owns the program (can be more than one)?
  - Master or Bachelor program, vocational training (or another type)?
  - List of tracks, courses, number of students (if available)
- Usage of other European products (for existing programs or for using existing courses for new programs).
  - Is your program aligned with European e-Competence Framework?
  - Is your program aligned with European Qualifications Framework?
- Are teaching-learning activities (TLA) and assessment methods (such as Problem-based Learning) explicitly considered in your program?

- Initial EDSF Reflection
  - Desk-based, assessing scientific and technical soundness, scope and relevance, readability, appearance and structure, usability, etc.
  - Expectations w.r.t. applying EDSF?
- EDSF Application
- CF-DS: Specification of matching competence groups, competences and learning outcomes
  - Balance of program w.r.t. Competence groups (what percentage of ECTS points per group)?
  - Are learning outcomes defined? Do they follow Bloom's Taxonomy?

- If learning outcomes defined, how do they align with EDISON learning outcomes in content and relative amount of ECTS points?
- Illustrative Examples (program, selected courses => complete mappings in xlsx)...
- Issues encountered... feedback (changes, additions, ...)

- BoK-DS: Specification of matching knowledge areas, topics
  - Illustrative Examples (program, selected courses => complete mappings in xlsx)...
  - Issues encountered... feedback (changes, additions, ...)

- Assignment of professional profiles targeted by program
  - What professional profile do you cover (of profiles defined in EDISON)?
  - Issues encountered... feedback (changes, additions, ...)

- MC-DS: Curricula Design and/or Evaluation

Depending on Use case A / B

- Gap analysis (remember to consider professional profiles, for existing programs or for using existing courses for new programs).
  - What are the gaps on the level of balance between competence groups?
  - What are the gaps on the level of learning outcomes?

- Summary and Outlook

Plans for addressing gaps/designing new program(s) so it/they complies with EDISON products.

## DS Programme Description Template

An excel spreadsheet with tables for DS programs and courses was provided.

| **Field Name** | | **Description** |
|---|---|---|
| *Title* | *Mandatory* | A meaningful short title |
| *Track Name* | *Optional* | Name of the track within the program |
| *Course List* | *Recommended* | URI to courses being part of the program, limited to the track if specified |
| *Organizer* | *Obligatory* | Institution, company, project organizing the program |
| *Type of Program* | *Mandatory* | Summer school, academic program, … |
| *Location* | *Mandatory* | A country and city (or full address) where the course takes place, unless online |
| *Start Date and Time* | *Mandatory* | The start date and time of the item (in ISO 8601 date format, preferably in UTC with time offset to local time zone). |
| *End Date and Time* | *Mandatory* | The end date and time of the item (in ISO 8601 date format, preferably in UTC with time offset to local time zone). |
| *URL* | *Mandatory* | Link to further information |
| *Contact* | *Mandatory* | Contact information of the responsible party (name, email or phone number) |
| *Language* | *Mandatory* | Language of instruction |
| *Level* | *Optional* | The level of studies following either Bologna or US approach |
| *Credit* | *Recommended* | Recommended for academic courses, including grading system |
| *Prerequisites* | *Recommended* | Required prior knowledge, preferably based on a BoK or taxonomy |
| *Target Audience* | *Optional* | E.g. “social scientists”, “biologists”, “data managers”, “policy makers in the UK”, or other |
| *Knowledge Areas* | *Recommended* | Knowledge areas covered by the program, based on the EDISON Body of Knowledge or Taxonomy |

| | | |
|---|---|---|
| *Competence Groups* | *Recommended* | Competence groups covered by the program, based on the EDISON Competence Framework |
| *Learning Outcomes* | *Recommended* | Including objectives, preferably based on the EDISON Competence Framework |
| *Professional Profiles* | *Recommended* | Professional Profiles addressed by program |
| *Description* | *Recommended* | E.g. The program will provide a strong basis in administrative, programing, and algorithm design aspects of data intensive systems. |
| *Registration Deadline* | *Optional* | The date and time of the item (in ISO 8601 date format, preferably in UTC with time offset to local time zone). |
| *Payment* | *Optional* | Use three letter currency symbols (in ISO 4217 format) and payment methods |

**Results from Champions / Use Case Partners A and B**

## University of Amsterdam (UvA)

- ## Summary of Data Science program and courses
  - Adam Belloum (track coordinator) Joined program between two universities (UvA and VUA), General information about program
    - Analyzing/improving an existing one
    - Computer Science
    - Track in Computer Science  Master program
    - List of tracks, courses
      Core Courses
      - Large-Scale Data Engineering (Peter Boncz) –Sep/Oct
      - Web Data Processing Systems  (Jacopo Urbani) –Nov/Dec
      - Information Visualization (Marcel Worring) - Jan
      - Data Mining Techniques (Mark Hoogendoorn) – Feb/Mar
      - Service Oriented Architecture and Cloud systems – Apr/May
      - Performance Engineering (Ana l. Varbanescu) –Apr/May
  - **Usage of other European products**:
    - Is your program aligned with European e-Competence Framework? No
    - Is your program aligned with European Qualifications Framework? No
  - **Are teaching-learning activities (TLA) and assessment methods (such as Problem-based Learning) explicitly considered in your program?** No
- ## Initial EDSF Reflection

  Applying EDSF will help to identify the target *DS science profiles* for which the graduates of UVA MSc DS program match in term of competencies and skills. At the same time using EDSF will help to identify gaps in the current offering on UVA MSc DS program, the program is it has been designed a year ago may have overlooked DS competences groups that are important for the targeted DS Science profiles.
- ## EDSF Application

- CF-DS: Specification of matching competence groups, competences and learning outcomes
  - **Balance of program w.r.t. Competence groups (what percentage of ECTS points per group)?**
    - Data Science Analytics competence group (**12**)
    - Data Science Engineering competence group (**18**)
    - Data Science Domain Knowledge (**6**)
    - Data Science Data Management (**0**)
    - Data Science Scientific Research Methods/ DSBPM Business Process (**36/6**)

  - **Are learning outcomes defined? Do they follow Bloom's Taxonomy?** Yes
    - UVA-VU data Science engineering track is part of the MSC program and mostly aim at assessment level
  - **If learning outcomes defined, how do they align with EDISON learning outcomes in content and relative amount of ECTS points.**
    - UVA-VU data Science engineering track is focusing on the Engineering aspects of the Data science profession, so in terms of ECTS points the ECTS the program is a bit biased toward engineering competences and skills

  - **Illustrative Examples**
    - All the 5 cores courses forming the DS engineering track have been mapped to the EDSF competence groups and Learning outcomes (cf. excel file)
  - **Issues encountered and feedback**
    - Mapping was quite forward and logical. IT shows strong point of the UVA-VU data Science engineering program (courses are complementing each other in terms of focus on DS competencies and Learning Objectives). But it showed also the weak points of the program like no course in the program is covering the DS data management competencies explicitly, while this is quite important for the Engineering aspects of Data science which is the focus of the program. One action that followed the alignment of the UvA DS engineering program with the EDSF is a start to preparing a new course that will be added as en elective for track and which will be dedicated to data management.

- BoK-DS: Specification of matching knowledge areas, topics
  - **Illustrative Examples**
    - Aligning UVA-VU data Science engineering courses with the EDSF in terms of competences and learning objectives was quite straightforward. Competences groups and learning objectives of courses where implicitly imbedded in the description of the courses and with the help of EDSF they could be identified and put in more clear and explicit manner.
  - **Issues encountered and feedback**
    - Aligning UVA-VU data Science engineering courses with the EDSF in terms of competences and learning objectives required some expertise in the field, the only obstacles we faced is the fact that competences groups and learning objectives are described sometimes at different level of granularity and used different terminology

- Assignment of professional profiles targeted by program
  - **What professional profile do you cover (of profiles defined in EDISON)?**
    Primary focus is DSP01-03, DSP04-09, as a secondary focus is DSP10-13

- MC-DS: Curricula Design and/or Evaluation

  Data science is an emerging field subject to frequents it is thus necessary to revise and adapt Data science curricula to keep them relevant to the data science job market. Because the EDISION methodology is taking into account Data Science professional profiles that are defined based on the ESCO occupation family profiles the EDISON data Science Framework can be easily adapted and thus form a good reference for institutions offering data science curricula and trainings to monitor the data science evolution and identify gaps in their Data Science offerings. Applying the EDSF to the UVA-VU data Science engineering program helped to identify a gap related to Data Management competences which are curial for many data science professional profiles but still not full covered in the current Data Science offerings.

  - **Gap analysis** (remember to consider professional profiles, for existing programs or for using existing courses for new programs).
    - **What are the gaps on the level of balance between competence groups?**
      - The Aligning UVA-VU data Science engineering track is not covering the DS data management competencies explicitly, while this is quite important for the Engineering aspects of Data science which is the focus of the program
    - **What are the gaps on the level of learning outcomes?**
      - Because the current Aligning UVA-VU data Science engineering track does not consider the competencies related to Data management, the associated Learning Objectives are not considered as well at least not explicitly.

# • Summary and Outlook

**Plans for addressing gaps/designing new program so it complies with EDISON products.**

As consequence of the alignment of UVA-VU data Science engineering track with the EDSF competences groups and Learning Objectives a new course that will address Data Management competences will be added to the of UVA-VU data Science engineering track. UVA staff members in concentration with the Research Data Alliance are defining the content of the course which be available as a series of trainings on *Research Data Management Literacy* and will be host at some of the RDA plenary meetings and thus available to All EDISON champion universities. The formal adoption of the topic Data Management in the UVA-VU data Science engineering track is expected for the next academic years).

| Field Name | Description |
|---|---|
| *Title* | **Program Big Data Engineering** |
| *Track Name* | Big Data Engineering |
| *Course List* | http://masters.vu.nl/en/programmes/computer-science-big-data-engineering/index.aspx |
| *Organizer* | Joined program between the UvA and VUA |
| *Type of Program* | 2 year MSc computer Science program |
| *Location* | Amsterdam, The Netherlands |
| *Start Date and Time* | 05.09.2017 |
| *End Date and Time* | Not Applicable |
| *URL* | http://masters.vu.nl/en/programmes/computer-science-big-data-engineering/index.aspx |
| *Contact* | Adam Belloum a.s.z.belloum@uva.nl, +31205257514 |
| *Language* | English |
| *Level* | MSC level |
| *Credit* | Track within the MSc computer Science program (30 ECTS out of 120 ECTS), could be combined with the research project 3 ECTS , literature study 6 ETCS, Industrial internship 6 ETCS |
| *Prerequisites* | BA in computer Science, BA in information Sciences plus programming experience, BA in Business Analytics with programming experience |
| *Target Audience* | Technical data managers, Data analysts working with Big Data infrastucture |
| *Knowledge Areas* | KAG2-DSENG: Data Science Engineering group including Software and infrastructure engineering |
| *Competence Groups* | Data Science Engineering |
| *Learning Outcomes* | learning Objectives LO3:01,02,03,04 ( see EDISON Data Science Model Curriculum (MC-DS): Approach and Initial version) |
| *Professional Profiles* | Professional data science (DS04-09) (see EDISON Data Science Model Curriculum (MC-DS): Approach and Initial version) |
| *Description* | E.g. The program will provide a strong basis developing Data Analytics programms using big data infrastucture, fundamentals of cloud systems, and . |
| *Registration* | 1 June for Dutch students. 1 April for EU/EEA and non-EU/EEA students |

| | |
|---|---|
| *Deadline* | |
| *Payment* | EUR (bank tansfer) |

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **Large-Scale Data Engineering** |
| *Name of Presenter(s)* | Prof. dr. P.A. Boncz, CWI Amsterdam |
| *Organizer* | CWI Amsterdam |
| *Type of Course* | Course |
| *Related Program* | http://masters.vu.nl/en/programmes/computer-science-big-data-engineering/index.aspx |
| *Location* | Amsterdam The Netherlands |
| *Start Date and Time* | 31/09/2017 |
| *End Date and Time* | 28.10.2017 |
| *URL* | http://event.cwi.nl/lsde/2016/index.shtml |
| *Contact* | prof. dr. P.A. Boncz |
| *Language* | English |
| *Level* | MSc level |
| *Credit* | 6 ECTS, Grading system: practical assignment 30%, presentation of literature study 20%, written reports of the literature study 40% |
| *Prerequisites* | KAG2-DSENG: AL, AR, SE |
| *Target Audience* | DSP01-03, DSP04-09,DSP014-16, |
| *Knowledge Areas* | KAG2-DSENG, |
| *Competence Groups* | DSENG01, DSENG04, |
| *Learning Outcomes* | LO3.01, LO3.02, LO3.03, LO3.04 |
| *Description* | The course further gives an overview of the infrastructures currently at the disposal of a broad public to address large scale data analysis |

| | |
|---|---|
| *Registration Deadline* | Enrolment in interdepartmental courses of VU, UvA and AUC first semester 2016-2017, June [Thursday 2 June 10 am till Monday 6 June 11 pm (2016)] |
| *Payment* | Part of MSc program, the registration for the program covers the course |

| | |
|---|---|
| ***Title*** | **Web Data Processing Systems** |
| *Name of Presenter(s)* | Jacopo Urbani, VUA |
| *Organizer* | Free University Amsterdam |
| *Type of Course* | course |
| *Related Program* | http://masters.vu.nl/en/programmes/computer-science-big-data-engineering/index.aspx |
| *Location* | Amsterdam The Netherlands |
| *Start Date and Time* | 31.10.2017 |
| *End Date and Time* | 23.12.2017 |
| *URL* | http://www.vu.nl/en/study-guide/2016-2017/master/m-r/parallel-and-distributed-computer-systems/index.aspx?view=module&origin=50049360&id=51248205 |
| *Contact* | Jacobo Urbani, email j.urbani@vu.nl |
| *Language* | English |
| *Level* | MSc level |
| *Credit* | 6 ECTS, grading (practical assignment + final exam) |
| *Prerequisites* | KAG2-DSENG: AL, AR |
| *Target Audience* | DSP04-09, DSP010-013, |
| *Knowledge Areas* | KAG2-DSDA, |
| *Competence Groups* | DSDA01, DSDA02, DSD04 |
| *Learning Outcomes* | LO1.01, LO1.02, LO1.04 |
| *Description* | The course introduces the student to the most advanced systems and techniques which deal with Web data, with a particular emphasis on scalability. Important classes of problems concern: storage and retrieval of Web data, efficient entity disambiguation, knowledge extraction, expressive ontological inference |

| | |
|---|---|
| *Registration Deadline* | Enrolment in interdepartmental courses of VU, UvA and AUC first semester 2016-2017, June [ Thursday 2 June 10 am till Monday 6 June 11 pm (2016)] |
| *Payment* | Part of MSc program, the registration for the program covers the course |

| | |
|---|---|
| ***Title*** | **Information Visualization** |
| *Name of Presenter(s)* | Macel Worring, UvA |
| *Organizer* | UVA |
| *Type of Course* | Course |
| *Related Program* | http://masters.vu.nl/en/programmes/computer-science-big-data-engineering/index.aspx |
| *Location* | Amsterdam The Netherlands |
| *Start Date and Time* | 02.01.2017 |
| *End Date and Time* | 31/02/2017 |
| *URL* | http://studiegids.uva.nl/xmlpages/page/2016-2017/zoek-vak/vak/23828 |
| *Contact* | Macel Worring, email:m.worring@uva.nl |
| *Language* | English |
| *Level* | MSc level |
| *Credit* | 6 ECTS, Grading system: Intermediate Report 10%, Final demo 10%, final report 30%, written exam 50% |
| *Prerequisites* | KAG2-DSENG: AL, SE |
| *Target Audience* | DSP04-09, DSP010-013, |
| *Knowledge Areas* | KAG2-DSDA, |
| *Competence Groups* | DSD06 |
| *Learning Outcomes* | LO1.06 |
| *Description* | The course covers the development of methodologies which support the process of gaining insight in large and complex datasets by a combination of data analysis, machine learning, and information visualization. |
| *Registration* | Enrolment in interdepartmental courses of VU, UvA and AUC second semester |

| | |
|---|---|
| *Deadline* | 2016-2017, December 2016 |
| *Payment* | Part of MSc program, the registration for the program covers the course |

| | |
|---|---|
| ***Title*** | **SOA-Cloud systems** |
| *Name of Presenter(s)* | Adam Belloum, UvA |
| *Organizer* | UvA |
| *Type of Course* | course |
| *Related Program* | http://masters.vu.nl/en/programmes/computer-science-big-data-engineering/index.aspx |
| *Location* | Amsterdam The Netherlands |
| *Start Date and Time* | 01.04.2017 |
| *End Date and Time* | 31.05.2017 |
| *URL* | http://studiegids.uva.nl/xmlpages/page/2016-2017/zoek-vak/vak/29398 |
| *Contact* | Adam Belloum, email: a.s.z.belloum@uva.nl |
| *Language* | English |
| *Level* | MSc level |
| *Credit* | 6 ECTS, Grading system: practical assignment 60%, presentation of literature study 20%, written reports of the literature study 20% |
| *Prerequisites* | KAG2-DSENG: AL, SE |
| *Target Audience* | DSP01-03, DSP04-09 |
| *Knowledge Areas* | KAG2-DSENG, |
| *Competence Groups* | DSENG01, DSENG02, DSENG05 |
| *Learning Outcomes* | LO3.01, LO3.02, LO3.05 |
| *Description* | The course introduces students to the principles of web services and cloud system. students will learn about the different paradigms of cloud systems (IaaS, PaaS, SaaS), and understand the mechanisms and technologies behind each mode to successfully harness cloud resources. |
| *Registration* | Enrolment in interdepartmental courses of VU, UvA and AUC second semester |

| | |
|---|---|
| *Deadline* | 2016-2017, December 2016 |
| *Payment* | Part of MSc program, the registration for the program covers the course |

| | |
|---|---|
| ***Title*** | **Performance Engineering** |
| *Name of Presenter(s)* | Ana Varbanescu |
| *Organizer* | UvA Amsterdam |
| *Type of Course* | Course |
| *Related Program* | http://masters.vu.nl/en/programmes/computer-science-big-data-engineering/index.aspx |
| *Location* | Amsterdam The Netherlands |
| *Start Date and Time* | 01.04.2017 |
| *End Date and Time* | 31.05.2017 |
| *URL* | http://studiegids.uva.nl/xmlpages/page/2016-2017/zoek-vak/vak/1544979 |
| *Contact* | Ana Varbanescu, email: a.l.varbanescu@uva.nl |
| *Language* | English |
| *Level* | MSc level |
| *Credit* | 6 ECTS, Grading system: practical assignment 25%, project 50%, exam 25% |
| *Prerequisites* | KAG2-DSENG: AL, AR |
| *Target Audience* | DSP01-03, DSP04-09 |
| *Knowledge Areas* | KAG2-DSENG, |
| *Competence Groups* | DSENG01, DSENG02, DSENG05 |
| *Learning Outcomes* | LO3.01, LO3.02, LO3.05 |
| *Description* | The course focuses on the modern aspects of performance engineering in the context of parallel algorithms and applications. Students learn to define and use performance metrics and measurement techniques, performance analysis, benchmarking and micro benchmarking, performance models (analytical and statistical), and performance prediction |
| *Registration* | Enrolment in interdepartmental courses of VU, UvA and AUC second semester |

| | |
|---|---|
| *Deadline* | 2016-2017, December 2016 |
| *Payment* | Part of MSc program, the registration for the program covers the course |

## University of Stavanger (UiS)

- Summary of Data Science program and courses
  - Tomasz Wiktorski, Lector at University of Stavanger
    - Designing a new Data Science Master program at Faculty of Science and Technology

  - **Usage of other European products:** No
- Initial EDSF Reflection

EDSF provides a good thinking framework to design data science program. In its current state it lack enough detail to provide specific support to design courses that are part of the program. The main element necessary to also provide that support is the connection between Competences/Learning Outcomes and BoK.

- **Expectations w.r.t. applying EDSF (Use Case description)?**

The goal is to design a new master program in Data Science, which should base to the biggest extent possible on existing courses. A difficulty is that based on internal policy UiS can only offer courses worth 10ECTS points, together with 30ECTS thesis, it means that student will only take 9 courses during the course of studies. An additional challenge is to allow for flexible design of the program to allow students with different backgrounds to take the program. The question is raised whether students missing programming competences could take such master's program.

- EDSF Application
- CF-DS: Specification of matching competence groups, competences... and learning outcomes

- **Balance of program w.r.t. Competence groups (what percentage of ECTS points per group)?**

On the program level the balance is achieved. Because it's a new program it aims to follow EDISON's approach closely. The problem appears during the transfer of learning outcomes from program level to course level.

- **Are learning outcomes defined? Do they follow Bloom's Taxonomy?**

On the course level learning outcomes are usually not defined properly. They do not explicitly consider Bloom's Taxonomy. It is an additional difficulty in machine courses to program.

- **If learning outcomes defined, how do they align with EDISON learning outcomes in content and relative amount of ECTS points.**

While learning outcomes are defined, because of problems mentioned in first two points, it is impossible to determine answer to this question.

- **Illustrative Examples**

Let's take course "Statistical modelling and simulation" as an example. It defined the following learning outcomes:

- be able to make and use statistical models for a number of technical, economic and scientific problems

- have knowledge of the strengths and limitations of some key techniques for statistical modelling

- be able to implement the models (R)

- carry out simulations of statistical models, analyse the results statistically, and

- to present results in a proper manner

- able to make assessments of uncertainty in the results

We can see that these learning outcomes cover the following EDISON LOs/Competences:

LO1.02 DSDA02 - Use appropriate statistical techniques on available data to deliver insights.

And partially

LO1.05 DSDA05 - Use different data analytics platforms to process complex data.

However, we can notice that learning outcomes are poorly defined. They don't follow Bloom's taxonomy and it is impossible to determine what level of competences they cover.

- **Issues encountered and feedback**

For many courses LOs are defined but it's difficult to decide whether they actually correspond to program's LOs. Course LO are often very specific and directed towards specific elements in BoK relevant for the course, while program LOs are general. Course LOs often don't follow Bloom's taxonomy what creates and additional challenge in matching.

### • BoK-DS: Specification of matching knowledge areas, topics

Currenty BoK-DS is not yet ready for application on the course level, mostly because of missing links between BoK and LOs. On the program level BoK is not as important, so it wasn't applied.

### • Assignment of professional profiles targeted by program

- **What professional profile do you cover (of profiles defined in EDISON)?**

DSP04 Data Scientist

DSP07 Data Science (Application) Programmer/Engineer

- **Issues encountered and feedback**

A general challenge is that depending on the department that drives the program there is a push to debalance the program towards the competence group closest to departments core competences. This is both due to academic and economic reasons.

### • MC-DS: Curricula Design and/or Evaluation

It is too early for this analysis.

## • Summary and Outlook

There are two general actions planned. There is a need to update LOs of individual courses so it is possible to better evaluate them in the context of the whole program. The second action is to contribute to linking LOs with BoK, how it could be done is still to be determined.

| Field Name | Description |
|---|---|
| ***Title*** | **Program Data Science** |
| *Track Name* | CS, STAT, DOM |
| *Course List* | N/A |

| | |
|---|---|
| *Organizer* | University of Stavanger |
| *Type of Program* | Academic |
| *Location* | Stavanger, Norway |
| *Start Date and Time* | 15.08.2018 |
| *End Date and Time* | 15.06.2020 |
| *URL* | not yet |
| *Contact* | tomasz.wiktorski@uis.no |
| *Language* | English |
| *Level* | Master |
| *Credit* | 120 ECTS |
| *Prerequisites* | N/A |
| *Target Audience* | All |
| *Knowledge Areas* | DSP04 |
| *Competence Groups* | N/A |
| *Learning Outcomes* | N/A |
| *Professional Profiles* | N/A |
| *Description* | This is a very early draft, an idea of a new master program was proposed to the Faculty of Science and Technology and should be hopefully approved in December, only then will actual work start. Meaning of tracks is different than usually. Track describe the competences of students entering the Data Science program. For instance, students with Computer Science BSc can enter program and they will have realtively fewer CS courses, but more courses in Statistics/Data Analysis and Domain Knowledge. The same respectively for students comming with Statistics BSc and Domain Knowledge BSc. |
| *Registration Deadline* | 1.02.2017 tentative |
| *Payment* | FREE |

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **Data-intensive Systems** |
| *Name of Presenter(s)* | Tomasz Wiktorski |
| *Organizer* | University of Stavanger |
| *Type of Course* | academic course |
| *Related Program* | Data Science |
| *Location* | Stavanger, Norway |
| *Start Date and Time* | 02.01.2019 |
| *End Date and Time* | 15.07.2019 |

| | |
|---|---|
| *URL* | http://www.uis.no/studies/study-courses/?categoryID=10648&parentcat=9835&code=DAT500_1&name=Data-Intensive+systems |
| *Contact* | tomasz.wiktorski@uis.no |
| *Language* | English |
| *Level* | master |
| *Credit* | 10 ECTS |
| *Prerequisites* | N/A |
| *Target Audience* | N/A |
| *Knowledge Areas* | N/A |
| *Competence Groups* | N/A |
| *Learning Outcomes* | Characterize Hadoop job tracker, task tracker, scheduling issues, communications, and resource management; describe elements of Hadoop ecosystem and identify their applicability; describe and compare RDBMS, data warehouse, unstructured big data, and keyed files, and show how to apply them to typical data processing problems; understand algorithmic complexity of the worst case, expected case, and best case running time, and the orders of complexity; apply the analysis to real life algorithms; design, construct, test, and benchmark a small data processing cluster (based on Hadoop); analyze real-life problems and propose suitable solutions; construct programs based directly on MapReduce paradigm for typical problems; construct programs based on high-level tools (for MapReduce paradigm) for typical problems; analyze influence of peak and sustained bandwidth rate on system performance; evaluate, communicate and defend a data-intensive solution w.r.t. relevant criteria |
| *Description* | The course will provide a strong basis in administrative, programing, and algorithm design aspects of data intensive systems. |
| *Registration Deadline* | N/A |
| *Payment* | N/A |

| | |
|---|---|
| ***Title*** | **Web Search and Data Mining** |
| *Name of Presenter(s)* | Krisztian Balog |
| *Organizer* | University of Stavanger |
| *Type of Course* | academic course |
| *Related Program* | Data Science |
| *Location* | Stavanger, Norway |
| *Start Date and Time* | 15.08.2018 |
| *End Date and Time* | 15.12.2018 |
| *URL* | http://www.uis.no/studies/study-courses/?code=DAT630_1 |
| *Contact* | krisztian.balog@uis.no |

| | |
|---|---|
| *Language* | English |
| *Level* | master |
| *Credit* | 10 ECTS |
| *Prerequisites* | N/A |
| *Target Audience* | N/A |
| *Knowledge Areas* | N/A |
| *Competence Groups* | N/A |
| *Learning Outcomes* | Theory and practice of data mining and information retrieval concepts, methods, and techniques. Process and prepare large-scale textual data collections for mining and retrieval.<br>Apply clustering, classification, and ranking methods to a range of information access problems.<br>Evaluate results and perform error analysis.<br>Understanding of the strengths and limitations of popular data mining and information retrieval techniques. Being able to identify promising business applications, participate in and lead such projects. |
| *Description* | The course offers an introduction to techniques and methods for processing, mining, and searching in massive (mostly textual) data collections. Topics range from statistics to machine learning, natural language processing, and information retrieval.<br><br>The course will study various applications and provide an opportunity for hands-on experimentation with state-of-the-art algorithms using existing software tools and data collections. |
| *Registration Deadline* | N/A |
| *Payment* | N/A |

| | |
|---|---|
| ***Title*** | **Statistical modeling and simulation** |
| *Name of Presenter(s)* | Bjørn Henrik Auestad, Dag Bjarne Tjøstheim, Tore Selland Kleppe |
| *Organizer* | University of Stavanger |
| *Type of Course* | academic course |
| *Related Program* | Data Science |
| *Location* | Stavanger, Norway |
| *Start Date and Time* | 15.08.2018 |
| *End Date and Time* | 15.12.2018 |
| *URL* | http://www.uis.no/studies/study-courses/?code=STA510_1 |
| *Contact* | Mandatory |
| *Language* | English |
| *Level* | master |
| *Credit* | 10 ECTS |

| | |
|---|---|
| *Prerequisites* | N/A |
| *Target Audience* | N/A |
| *Knowledge Areas* | N/A |
| *Competence Groups* | N/A |
| *Learning Outcomes* | be able to make and use statistical models for a number of technical, economic and scientific problems; have knowledge of the strengths and limitations of some key techniques for statistical modeling; be able to implement the models (R); carry out simulations of statistical models, analyze the results statistically, and to present results in a proper manner; able to make assessments of uncertainty in the results |
| *Description* | This course provides a foundation for problem solving in technology, economics and science using statistical modeling, simulation and analysis. The exam is done with PC on PC-lab, parts on PCs and parts on plain paper. |
| *Registration Deadline* | N/A |
| *Payment* | N/A |

## Goethe University Frankfurt (GOE)

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **Big Data Technology** |
| *Name of Presenter(s)* | Prof. Dott. Ing. Roberto V. Zicari, Todor Ivanov and Kim Hee - Goethe University |
| *Organizer* | GOETHE University - Big Data Lab |
| *Type of Course* | Academic course |
| *Related Program* | N/A |
| *Location* | Frankfurt am Main(Germany) |
| *Start Date and Time* | 23.10.2017 |
| *End Date and Time* | 05.02.2018 |
| *URL* | http://www.bigdata.uni-frankfurt.de/big-data-technologies-ws-2017/ |
| *Contact* | Roberto Zicari, zicari@dbis.cs.uni-frankfurt.de |
| *Language* | English and German |
| *Level* | Master Students in Computer Science, Bio informatics and Business informatics |
| *Credit* | 8 CPs |
| *Prerequisites* | programming skills, knowledge of Python, algorithms and data structures |
| *Target Audience* | N/A |
| *Knowledge Areas* | DSA, DSE |
| *Competence Groups* | DSENG03, DSENG04, DSENG06; DSDM02, DSDM05; |
| *Learning* | N/A |

| | |
|---|---|
| *Outcomes* | |
| *Description* | N/A |
| *Registration Deadline* | 19 october 2017 |
| *Payment* | N/A |

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **Introduction to Data Science** |
| *Name of Presenter(s)* | Dr. Jochen L. Leidner (Thomson Reuters) and Kim Hee Goethe University) |
| *Organizer* | Goethe University |
| *Type of Course* | Academic course |
| *Related Program* | N/A |
| *Location* | Frankfurt am Main (Germany) |
| *Start Date and Time* | 29 May 2017 |
| *End Date and Time* | 15 sept 2017 |
| *URL* | http://www.bigdata.uni-frankfurt.de/introduction-data-science/ |
| *Contact* | Hee@dbis.cs.uni-frankfurt.de |
| *Language* | English and German |
| *Level* | Master Students in Computer Science, Bio informatics and Business informatics |
| *Credit* | 5 CPs |
| *Prerequisites* | programming skills, knowledge of Python, algorithms and data structures |
| *Target Audience* | N/A |
| *Knowledge Areas* | DSA, DSE |
| *Competence Groups* | DSENG03, DSENG04, DSENG06; DSDM02, DSDM05; |
| *Learning Outcomes* | N/A |
| *Description* | The goal of this compact course is to give participants a first gentle introduction and solid conceptual grounding in what has been called 'data science', i.e. experimental work that is data-driven and empirical. The focus is on methodology, defining an experimental protocol, devising hypotheses, thinking about measuring success, but also on more practical approaches like basic machine learning methods (both supervised and unsupervised) and natural language processing approaches (like part-of-speech tagging, named entity recognition/classification/resolution, and parsing) and the introduction to popular tools. The course also demonstrates some practical applications of the techniques shown, and deepens the students' skills via practical exercises.<br><br>The lecture is delivered over 4 weeks of calendar time and consists of 2 three-day blocks of 3 hours of lectures followed by 2 days of 2.5 hours of exercises/tutorials each). It targets Master's level students. By the end of the course, participants will be able to analyze data-sets, and to create their own predictive classifieds and visualizations. |

| | |
|---|---|
| *Registration Deadline* | 19 Oct 2017 |
| *Payment* | N/A |

## University of Bedfortshire (BEDS)

- ## Summary of Data Science program and courses

From the academic year 2017/18 on, the University of Bedfordshire will offer two undergraduate Data Science programmes: BSc (Hons) Data and Information Systems and BSc (Hons) Computing and Data Science. Both programmes emphasise the information systems/computer science view on Data Science whilst still including mathematical aspects.

- Contact information
  Ingo Frommholz, School of Computer Science and Technology, Park Square, Luton LU1 3JU, UK
- General information about program
  - The School of Computer Science has been running a BSc (Hons) Data Science programme in its Milton Keynes campus. However, this programme has not recruited well, likely due to a 'labelling problem'. To attract more students to Data Science, the following actions were performed:
    - Rename BSc (Hons) Data Science to BSc (Hons) Computing and Data Science
    - Replace the old BSc (Hons) Information Systems programme with BSc (Hons) Information and Data Systems

    Both new programmes are currently subject to final university approval and it is expected they will be offered for the next academic year 2017/18.

    The School of Computer Science and Technologies considers the creation of new MSc-level programmes on Data Science and on Information Retrieval, respectively.

  - School of Computer Science and Technology is responsible for these Bachelor programs
  - List of tracks, courses, number of students (if available)
    Courses: see spreadsheet. Student numbers not available so far, expected are less than 10 initially.

- Usage of other European products (for existing programs or for using existing courses for new programs).
  - Is your program aligned with European e-Competence Framework? **no**
  - Is your program aligned with European Qualifications Framework? **no**

- Are teaching-learning activities (TLA) and assessment methods (such as Problem-based Learning) explicitly considered in your program?

  Students are assessed in a variety of ways, including coursework, group and individual projects, portfolios, essays, presentations or exams. Students will also produce software artefacts in the area of their specialism. Constant feedback and advice from a supervisory or module team will be provided.

## • Initial EDSF Reflection

We considered EDISON for the design and approval of our two newly created programmes mentioned above as follows:

- CF-DS and BoK-DS were used as 'benchmarks' for our programmes and as guidance for creating the curricula (e.g., by embedding our modules in competence groups to make sure we cover all relevant aspects of Data Science)
- The EDISON outcome gave use reassurance that our programmes are in line with the demands of the Data Science market
- The fact that through EDISON we are well connected with Europe-wide Data Science experts was well received by the panel in the approval event of our programmes.

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **BSc (Hons) Information and Data Systems** |
| *Track Name* | |
| *Course List* | Introduction to Software Development<br>Principles of Programming<br>Computer Systems Structure<br>Mathematics of Data<br>Operational Information Security Management<br>Distributed Data Management and Semantics<br>Decision Support Systems and Data Mining<br>Systems Development and Modern Database Practices<br>Professional Practice Year (Computer Science and Technology, optional)<br>Research Methodologies and Emerging Technologies<br>Data Engineering, Presentation and Retrieval<br>Social and Professional Project Management<br>Undergraduate Project |
| *Organizer* | School of Computer Science and Technology, Faculty for Creative Arts and Technology, University of Bedfordshire |
| *Type of Program* | Academic program |
| *Location* | Luton, UK |
| *Start Date and Time* | 01.10.2017 |
| *End Date and Time* | 30.06.2020 |
| *URL* | tba |
| *Contact* | Ingo Frommholz <ingo.frommholz@beds.ac.uk> |
| *Language* | English |
| *Level* | England FHEQ Level 6 (see https://www.naric.org.uk/Europass/documents/ds_chart.pdf) |
| *Credit* | 360 |
| *Prerequisites* | A-levels or equivalent. At least one grade of at least 'B' in Mathematics or Physics is required (or equivalent qualification for international applicants) |
| *Target Audience* | Data Science Professionals, Database and Network Professionals, Data Science Technology Professionals |

| | |
|---|---|
| *Knowledge Areas* | Data Science Analytics/Mathematics of computing/Mathematical software; Data Science Analytics/Computing methodologies/Articifical Intelligence/Search metholdologies; Data Science Analytics/Computing Methodologies/Machine learning; Data Science Analytics/Information Systems/Information Systems Applications;Data Science Data Management/Data management systems; Data Science Engineering/Software and its engierring/Software creation and management; Data Science Engineering/Software and its engierring/Software notations and tools; Domain knowledge/Applied computing/Document management and text processing<br><br>DS-BoK Release 2: KU1.01.01, KU1.01.02, KU1.01.14, KU1.02.02, KU1.02.03, KU1.02.08, KU1.04.02, KU1.04.05, KU2.01.01, KU2.01.07, KU2.01.02, KU2.02.02, KU2.02.08, KU2.06.03, KU2.06.05, KU2.07.01, KU2.07.05, KU3.01.02, KU3.01.04, KU3.01.07, KU3.02.02, KU3.02.04, KU3.03.05, KU3.06.01, KU3.06.02, KU3.06.04, KU4.01.01, KU4.01.02, KU4.01.04, KU4.01.06, KA04.01 DSRMP.02/PM Project Management |
| *Competence Groups* | Domain Expertise, Research, Data Analytics, Analytic Systems, Algorithms, Engineering Competencies, Scientific Methods |
| *Learning Outcomes* | • LO1 Research, analyse and evaluate technologies and organisational problems in a range of contexts and to choose and implement appropriate solutions<br>• LO2 Demonstrate skills that allow you to conceptualise and apply formal and informal creative thinking techniques towards the development and implementation of relevant data and information systems into the real world<br>• LO3 Demonstrate skills in project management in relation to the delivery of projects within the constraints of client critical success factors<br>• LO4 Express, interpret and critically evaluate issues concerning the law and professional ethics in the context of information systems<br>• LO5 Demonstrate investigative skills in the area of information and data systems through completion of substantial assignments, reports, presentations and case studies<br>• LO6 Apply skills in a rational argument, objective interpretation of evidence, judgement and decision making towards the planning, analysis and successful development of information systems within complex organisational environments<br>• LO7 Understand and extend existing information systems and data science concepts, theories and practices following professional conventions and standards<br>• LO8 Critically describe and evaluate future trends in information and data systems development<br>• LO9 Work effectively within a systems development team and be able to explain the conditions necessary for successful team working |
| *Professional Profiles* | |
| *Description* | The course will provide you with academic and technical skills to analyse, interpret and make sense of how to handle the huge amount of ubiquitous data created daily. It will enable you to meet the demands of tomorrow's information society and draws from our expert staff in Information Systems, Data Science and Computing. |
| *Registration Deadline* | |
| *Payment* | |

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **BSc (Hons) Computing and Data Science** |
| *Track Name* | |

| | |
|---|---|
| *Course List* | Introduction to Software Development<br>Principles of Programming<br>Computer Systems Structure<br>Mathematics of Data<br>Operational Information Security Management<br>Distributed Data Management and Semantics<br>Decision Support Systems and Data Mining<br>Concepts and Technologies of Artificial Intelligence<br>Professional Practice Year (Computer Science and Technology, optional)<br>Research Methodologies and Emerging Technologies<br>Data Engineering, Presentation and Retrieval<br>Social and Professional Project Management<br>Undergraduate Project |
| *Organizer* | School of Computer Science and Technology, Faculty for Creative Arts and Technology, University of Bedfordshire |
| *Type of Program* | Academic program |
| *Location* | Milton Keynes, UK |
| *Start Date and Time* | 01.10.2017 |
| *End Date and Time* | 30.06.2020 |
| *URL* | tba |
| *Contact* | Ingo Frommholz <ingo.frommholz@beds.ac.uk> |
| *Language* | English |
| *Level* | England FHEQ Level 6 (see https://www.naric.org.uk/Europass/documents/ds_chart.pdf) |
| *Credit* | 360 |
| *Prerequisites* | A-levels or equivalent. At least one grade of at least 'B' in Mathematics or Physics is required (or equivalent qualification for international applicants) |
| *Target Audience* | Data Science Professionals, Database and Network Professionals, Data Science Technology Professionals |
| *Knowledge Areas* | Data Science Analytics/Mathematics of computing/Mathematical software; Data Science Analytics/Computing methodologies/Articifical Intelligence/Search metholdologies; Data Science Analytics/Computing Methodologies/Machine learning; Data Science Analytics/Information Systems/Information Systems Applications;Data Science Data Management/Data management systems; Data Science Engineering/Software and its engineering/Software creation and management; Data Science Engineering/Software and its engierring/Software notations and tools; Domain knowledge/Applied computing/Document management and text processing<br><br>DS-BoK Release 2: KU1.01.01, KU1.01.02, KU1.01.14, KU1.02.02, KU1.02.03, KU1.02.08, KU1.04.02, KU1.04.05, KU2.01.02, KU2.02.02, KU2.02.08, KU2.06.03, KU2.06.05, KU2.07.01, KU2.07.05, KU3.01.02, KU3.01.04, KU3.01.07, KU3.02.02, KU3.02.04, KU3.03.05, KU3.06.01, KU3.06.02, KU3.06.04, KU4.01.01, KU4.01.02, KU4.01.04, KU4.01.06, KA04.01 DSRMP.02/PM Project Management |
| *Competence Groups* | Domain Expertise, Research, Data Analytics, Analytic Systems, Algorithms, Engineering Competencies, Scientific Methods |

| | |
|---|---|
| *Learning Outcomes* | • LO1: Demonstrate a thorough understanding of the relevant areas of computing and data science such as: data storage and retrieval, data presentation, data security, data mining and data management; as well as the societal context in which data is used.<br>• LO2: Critically apply and reflect upon data analysis and modelling, presentation and other essential concepts along with the fundamental principles of data retrieval in the context of industry relevant case studies, showing judgement in using the right analysis and presentation tools and techniques.<br>• LO3: Professionally conduct a major project, using problem-solving and evaluation skills relevant to Data Science that addresses all aspects of a software development cycle including design, implementation and documentation in a coherent and consistent way.<br>• LO4: Demonstrate transferable skills and an ability to work under guidance and as a team member.<br>• LO5: Identify how the methods, tools and techniques of data science are applied within a professional, legal and ethical framework and develop an understanding of the need of continuing professional development.<br>• LO6: Critically discuss societal applications of Data Science.<br>In order to qualify for the award of BSc (Hons) Computing and Data Science (with Professional Practice year) students will need to meet all of the outcomes above and:<br>• LO7: Demonstrate knowledge and analytical understanding of professional practice by successfully completing an approved period of approved work place practice. |
| *Professional Profiles* | |
| *Description* | The course will provide you with academic and technical skills to analyse, interpret and make sense of the huge amount of ubiquitous data created daily. It will enable you to meet the demands of tomorrow's information society. |
| *Registration Deadline* | |
| *Payment* | |

## Lodz University of Technology (TUL)

- Summary of Data Science program and courses
  - Lodz University of Technology (TUL), Institute of Electronics
    Contact person: Aleksandra Królak, aleksandra.krolak@p.lodz.pl
    - Analyzing/improving existing Bachelor program at Institute of Electronics, Department of Medical Electronics in cooperation with Institute of Materials Science and Engineering, Department of Microelectronics and Computer Science and Department of Material and Commodity Sciences and Textile Metrology
    - List of tracks, courses
      Single track: Biomedical Engineering
      List of courses:
      http://programy.p.lodz.pl/kierunekSiatka.jsp?l=en&w=Biomedical%20Engineering&p=5493&stopien=first-cycle%20programme&tryb=full-time
  - **Usage of other European products (for existing programs or for using existing courses for new programs).**
    - **Is your program aligned with European e-Competence Framework?**
      partly
    - **Is your program aligned with European Qualifications Framework?**
      partly
  - **Are teaching-learning activities (TLA) and assessment methods (such as Problem-based Learning) explicitly considered in your program?**

Activity types included in the program are mainly Interactivity focus (group assignments), Critical thinking (response to an assigned text) and Problem solving (group projects including design and implementation of an algorithm, device, etc.). In the program for 4th semester Team Project is planned as a 10 ECTS course that is realized in the Problem Based Learning form.

- ## Initial EDSF Reflection
  - The EDSF documents should be readable and understandable for people working on development of academic study programs. Available releases are well structured and contain a lot of important, useful and well illustrated information about DS framework, BoK and professional profiles. Some of the introductions are in my opinion a bit too extensive, however allow a reader to take a deeper look at the related approaches. For persons with initial knowledge about the content of EDSF documents it is easy to get needed information from the text. Persons who get the documents for the first time need to put relatively much effort to understand everything and use the documents efficiently.
  - **Expectations w.r.t. applying EDSF (Use Case description)?**
    EDSF documents include competence framework, body of knowledge, model curriculum and professional profiles description. Information contained in these documents should give clear guidelines how to develop new study program or modify existing one for student education at Bachelor and/or Master level in Data Science related discipline. In case of Biomedical Engineering study program one of the possible tracks may be Biomedical Data Analytics. EDSF documents should be the source of information about proportions of courses from different fields according to Bloom's Taxonomy for Data Science and should facilitate completing required at Polish universities study program descriptions of: initial requirements, learning outcomes it three groups (knowledge, skills, social competences) and occupational profiles of graduates.

- ## EDSF Application
- ### CF-DS: Specification of matching competence groups, competences... and learning outcomes
  - **Balance of program w.r.t. Competence groups (what percentage of ECTS points per group)?**
    The courses can be divided into 5 categories where 3 of them refer to the DS competence groups: General courses (foreign language, physical education) 10.5%, basic technical courses (Chemistry, Physics, Mechanical Engineering, etc.) 11.4%, DS Analytics 25.2%, DS Engineering 25.7%, Domain Knowledge 27.2%.
  - **Are learning outcomes defined? Do they follow Bloom's Taxonomy?**
    Learning outcomes are defined for each course in three groups: knowledge, skills, social competences, what is required by Polish Ministry of Science and Higher Education. LOs with their description are selected from lists developed by particular universities, in this case by Lodz University of Technology, and are not aligned with Bloom's Taxonomy
  - **If learning outcomes defined, how do they align with EDISON learning outcomes in content and relative amount of ECTS points.**
    LO1-DA – 11% (22ECTS)
    LO2-DM – 0% (1ECTS)
    LO3-ENG – 13% (27ECTS)
    LO4-RM – 0% (0ECTS)
    LO5-BPM – 5% (10ECTS)
  - **Illustrative Examples**

| Competence group | TUL Biomedical Engineering program courses |
|---|---|
| DS Analytics | Information Technology, Algorithms and Data Structures, Mathematics 1 and 2, Image Processing and Computer Graphics, Medical Informatics |
| DS Engineering | Signal Processing, Control Systems, Sensors, Medical Imaging |
| Domain Knowledge | Biomaterials 1 and 2, Medical Electronics, Implants and Artificial Organs, Microsystems in Medical Applications |

- **Issues encountered and feedback**
  In my opinion students at Lodz University of Technology have too much basic technical courses in the first and second semester. They are provided with a lot of theoretical knowledge that is not related to practical examples. These courses are taught in large groups for different study programs. A good idea would be to add different projects or laboratories matching particular study program that would show usefulness of presented theory in future professional life.

## BoK-DS: Specification of matching knowledge areas, topics

- **Illustrative Examples**

| Knowledge area groups | TUL Biomedical Engineering program courses |
|---|---|
| KAG1-DSDA | None |
| KAG2-DSENG | Algorithms and Data Structures, Signal Processing, Medical Informatics |
| KAG3-DSDM | Information Technology, Medical Informatics |
| KAG4-DSRM | Signal Processing, Medical Informatics |
| KAG5-DSBPM | None |

- **Issues encountered and feedback**
  TUL Biomedical Engineering program is more focused on engineering than on DS related fields. However in todays world in many “biomedical engineering” fields, such as biomedical signal acquisition and analysis, biomedical image processing or medical devices design it is necessary to have knowledge about methods of analysis, storing, securing and visualization of large amounts of data. The current program should be modified so that all above mentioned topics are covered.

## Assignment of professional profiles targeted by program

- **What professional profile do you cover (of profiles defined in EDISON)?**
  DSP04, DSP05, DSP07, DSP08
- **Issues encountered and feedback:**
  The program should include courses related to Data Bases what is necessary for any manipulation of large amounts of data. Graduate students report the need of introducing teaching of more programming languages in the study program.

## MC-DS: Curricula Design and/or Evaluation

- **Gap analysis (remember to consider professional profiles, for existing programs or for using existing courses for new programs).**

- **What are the gaps on the level of balance between competence groups?**
  The balance between 3 main EDSF competence groups is kept (around 25%-26%-27%). However the courses in TUL Biomedical Engineering program related to DS Analytics and DS Engineering do not cover all needed fields indicated in EDSF, such as Statistics, Research methods, Databases, Data visualization, Data security.
- **What are the gaps on the level of learning outcomes?**
  On the level of Learning Outcomes there are huge gaps with relation to MC-DS defined in EDISON. Due to the fact that the LOs defined in the program of TUL Biomedical Engineering are done in accordance with Polish standards they coincide in about 10% only

## • Summary and Outlook

It is necessary to introduce large changes in the program of Biomedical Engineering (TUL) so that it matches all requirements presented in EDISON documents. All learning outcomes and professional profiles must be rearrange according to EDISON guidelines. More courses related do data analysis, data visualization, data transmission and security must be introduced. Necessary is to include Database course in the program. As a member of the Teaching Committee for Biomedical Engineering program I will suggest such changes to the rest of the members.

| Field Name | Description |
|---|---|
| ***Title*** | **Biomedical Engineering** |
| *Track Name* | Biomedical Engineering |
| *Course List* | http://programy.p.lodz.pl/kierunekSiatka.jsp?l=en&w=Biomedical%20Engineering&p=4291&stopien=first-cycle%20programme&tryb=full-time |
| *Organizer* | Lodz University of Technology |
| *Type of Program* | Academic Program |
| *Location* | Łódź, Poland |
| *Start Date and Time* | 01.10.2017 08:00 |
| *End Date and Time* | 20.02.2020 23:59 |
| *URL* | http://programy.p.lodz.pl/?l=en&s=karta-opisu-programu-ksztalcenia&pk=Biomedical%20Engineering&pkId=221 |
| *Contact* | Aleksandra Królak aleksandra.krolak@p.lodz.pl |
| *Language* | English |
| *Level* | BSc |
| *Credit* | 210 ECTS (30ECTS per semester) |
| *Prerequisites* | Mathematics, Physics, Biology, Chemistry |
| *Target Audience* | Engineers, Data Sceintists, biomaterial engineer |
| *Knowledge Areas* | |
| *Competence Groups* | |
| *Learning Outcomes* | http://programy.p.lodz.pl/?l=en&s=efekty-ksztalcenia-wiedza&pk=Biomedical%20Engineering&pkId=221 http://programy.p.lodz.pl/?l=en&s=efekty-ksztalcenia- |

| | |
|---|---|
| | umiejetnosci&pk=Biomedical%20Engineering&pkId=221<br>http://programy.p.lodz.pl/?l=en&s=efekty-ksztalcenia-kompetencje-spoleczne&pk=Biomedical%20Engineering&pkId=221 |
| *Professional Profiles* | Graduates who hold an engineering degree in Biomedical Engineering have the ability to use modern equipment, diagnostic and therapeutic methods that implement modern electronic and teleinformation techniques. They are trained and qualified in terms of:<br>- co-operation with medical doctors on<br>+ integration, operation, use and maintenance of medical equipment<br>+ the use of diagnostic and therapeutic systems<br>- manufacturing and design of medical, diagnostic and therapeutic devices<br>- implementation of projects related to biomedical engineering.<br>Graduates may be employed in hospitals, clinics, and other health system units, companies that manufacture equipment and medical devices, departments of accreditation and attestation of instruments and medical devices; companies that design and implement medical devices, research institutions, consulting services and medical administration.<br>Graduates are familiar with a foreign language at the B2 level, and have the ability to use the technical language in the field of biomedical engineering. Graduates are prepared to enroll on second level studies and work in national and international interdisciplinary teams. |
| *Description* | The program provides strong basis in engineering, medical equipment design and biomaterial technologies. Students acquire knowledge from the fild of material engineering, electronics, textronics and computer technologies for biomedical applications |
| *Registration Deadline* | end of July |
| *Payment* | no payment |

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **Information Technology** |
| *Organizer* | Lodz University of Technology |
| *Type of Course* | academic course |
| *Related Program* | http://programy.p.lodz.pl/przedmiot.jsp?l=en&idPrzedmiotu=146771&s=1&j=0&w=Biomedical%20Engineering |
| *Location* | Lodz, Poland, Lodz University of Technology |
| *Start Date and Time* | 01.10.2017 |
| *End Date and Time* | 20.02.2018 |
| *URL* | http://programy.p.lodz.pl/przedmiot.jsp?l=en&idPrzedmiotu=146771&s=1&j=0&w=Biomedical%20Engineering |
| *Contact* | aleksandra.krolak@p.lodz.pl |
| *Language* | English |
| *Level* | BSc |
| *Credit* | 3 ECTS |
| *Prerequisites* | Basic skills in computer operation |
| *Target Audience* | data managers |
| *Knowledge* | KAG3 |

| | |
|---|---|
| *Areas* | |
| *Competence Groups* | DS Analytics |
| *Learning Outcomes* | LO01 |
| *Description* | Modern computer systems and information technology are presented for the students. The lecture includes computer hardware and system software. Data transmission, different communication channels and their parameters, development of local area networks have been discussed. Laboratory classes comprise tasks concerning networking software, applications from MS-Office package and software for graphical processing. |
| *Registration Deadline* | N/A |
| *Payment* | no payment |

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **Algorithms and Data Structures** |
| *Organizer* | Lodz University of Technology |
| *Type of Course* | academic course |
| *Related Program* | http://programy.p.lodz.pl/przedmiot.jsp?l=en&idPrzedmiotu=158570&s=1&j=0&w=Biomedical%20Engineering |
| *Location* | Lodz, Poland, Lodz University of Technology |
| *Start Date and Time* | 01.10.2017 |
| *End Date and Time* | 20.02.2018 |
| *URL* | http://programy.p.lodz.pl/przedmiot.jsp?l=en&idPrzedmiotu=158570&s=1&j=0&w=Biomedical%20Engineering |
| *Contact* | aleksandra.krolak@p.lodz.pl |
| *Language* | English |
| *Level* | BSc |
| *Credit* | 2 ECTS |
| *Prerequisites* | none |
| *Target Audience* | N/A |
| *Knowledge Areas* | KAG2 |
| *Competence Groups* | DS Analytics |
| *Learning Outcomes* | LO01 |
| *Description* | The course will provide a strong basis in programing and algorithm design |
| *Registration Deadline* | N/A |
| *Payment* | no payment |

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **Signal Processing** |
| *Organizer* | Lodz University of Technology |

| | |
|---|---|
| *Type of Course* | academic course |
| *Related Program* | http://programy.p.lodz.pl/przedmiot.jsp?l=en&idPrzedmiotu=150713&s=3&j=0&w=Biomedical%20Engineering |
| *Location* | Lodz, Poland, Lodz University of Technology |
| *Start Date and Time* | 01.10.2018 |
| *End Date and Time* | 20.02.2019 |
| *URL* | http://programy.p.lodz.pl/przedmiot.jsp?l=en&idPrzedmiotu=150713&s=3&j=0&w=Biomedical%20Engineering |
| *Contact* | aleksandra.krolak@p.lodz.pl |
| *Language* | English |
| *Level* | BSc |
| *Credit* | 5 ECTS |
| *Prerequisites* | The student knows fundamentals of mathematical analysis and matrix algebra |
| *Target Audience* | engineers, data scientists |
| *Knowledge Areas* | KAG2, KAG4 |
| *Competence Groups* | DS Engineering |
| *Learning Outcomes* | LO01, LO03, LO04, LO05 |
| *Description* | Teaching the studnets how to apply digital techniques for recording, processing and analysis of signals, with special emphasis on biomedical signals |
| *Registration Deadline* | N/A |
| *Payment* | no payment |

| **Field Name** | **Description** |
|---|---|
| ***Title*** | **Medical Informatics** |
| *Organizer* | Lodz University of Technology |
| *Type of Course* | academic course |
| *Related Program* | http://programy.p.lodz.pl/przedmiot.jsp?l=en&idPrzedmiotu=150726&s=5&j=0&w=Biomedical%20Engineering |
| *Location* | Lodz, Poland, Lodz University of Technology |
| *Start Date and Time* | 01.10.2019 |
| *End Date and Time* | 20.02.2020 |
| *URL* | http://programy.p.lodz.pl/przedmiot.jsp?l=en&idPrzedmiotu=150726&s=5&j=0&w=Biomedical%20Engineering |
| *Contact* | aleksandra.krolak@p.lodz.pl |
| *Language* | English |
| *Level* | BSc |
| *Credit* | 3 ECTS |
| *Prerequisites* | none |
| *Target* | data managers, biomedical engineers |

| | |
|---|---|
| *Audience* | |
| *Knowledge Areas* | KAG2, KAG3, KAG4 |
| *Competence Groups* | DS Engineering |
| *Learning Outcomes* | LO01, LO03, LO04, LO05 |
| *Description* | The knowledge transmission in the field Medical Informatics. |
| *Registration Deadline* | N/A |
| *Payment* | no payment |

## University of Perugia (UoP)

| Field Name | Description |
|---|---|
| ***Title*** | **Master Degree in Data Science** |
| *Track Name* | N/A |
| *Course List* | http://www.unipg.it/en/courses/bachelor-master-degrees/course-catalogue-2016-17?controller=corso&layout=default&corso=666&tab=DID#ancora-curriculum-1464 |
| *Organizer* | University of Perugia - Dept of Engineering |
| *Type of Program* | Academic Program |
| *Location* | Perugia (Italy) |
| *Start Date and Time* | 1 November 2017 |
| *End Date and Time* | 28 February 2019 |
| *URL* | http://www.unipg.it/en/courses/bachelor-master-degrees/course-catalogue-2016-17?controller=corso&layout=default&anno=2016&corso=666 |
| *Contact* | Paolo Valigi - valigi@unipg.it |
| *Language* | Italian and English |
| *Level* | N/A |
| *Credit* | 120 ECTS |
| *Prerequisites* | An essential requirement /prerequisite to access to the master degree program is the possession of a bachelor's degree, or another qualification obtained abroad and recognized as valid, in one of the following classes:<br><br>L-08 Degrees in Information Engineering<br><br>L-09 Degrees in Industrial Engineering<br><br>L-30 Degrees in Physical Sciences and Technologies L-31 Degrees in Computer Science<br><br>L-35 Degrees in Mathematical Sciences<br><br>L-41 degrees in Statistics<br><br>or corresponding classes according to the dm 509/99: |

| | |
|---|---|
| | 09 Degrees in Information Engineering<br><br>10 Degrees in Industrial Engineering<br><br>26 Degrees in Computer Science<br><br>25 Degrees in Physical Sciences and Technologies<br><br>32 Mathematics<br><br>37 Statistics<br><br>In addition to bachelor's degree in one of the above classes, access requires possession of curricular requirements and adequate personal preparation.<br><br>In particular:<br><br>- The curricular requirements will be evaluated in terms of possession of a sufficient number of credits in the areas related to the basic and the characterizing activities of the L8 class with particular reference to the areas Electronics Engineering, Computer Engineering and Telecommunications Engineering.<br><br>- Personal preparation is assessed by setting defined thresholds based on the first-level degree mark, or on the grade obtained in various tests, or a combination thereof.<br><br>For candidates who do not pass the required thresholds, the degree program prepares a test of personal preparation. The procedures for verification of the curricular requirements and personal preparation are set out in the Regulations of the Degree program. |
| *Target Audience* | Bachelor Students |
| *Knowledge Areas* | DSA, DSE, DSBP |
| *Competence Groups* | DSDA01, DSDA02, DSDA04, DSDS06; DSENG01, DSENG02, DSENG03, DSENG06; DSDM02; DSDM05; DSBA01, DSBA03 |
| *Learning Outcomes* | The Master of Science in Computer Engineering and Robotics aims to train engineers in computing and automation, and to provide its students the references and the ability to contribute to innovation and to stay continuously updated with new technologies.<br>As specific objectives, the course aims to train experts in areas of great interest and development, such as: the processing and extraction of large amounts of data information; tools and technologies for robotics and industrial applications, in civil and service industry context. |
| *Professional Profiles* | DSP04, DSP05, DSP06, DSP07, DSP08, DSP17 |

| | |
|---|---|
| *Description* | The curriculum is as follows:<br><br>- In the first year course students will acquire advanced knowledge on digital signal processing and statistical analysis in the engineering field. They'll study the models and the main algorithmic techniques for machine learning and data mining, which represent an important knowledge for the modern computer science engineer's training.<br><br>- In a second step (which involves both the first and the second year), students focus on specialized disciplines more oriented to information technology (such as software engineering, computer security, distributed programming and computational complexity), or on specialized disciplines more oriented to the field of robotics (such as control, automation, and embedded electronic systems). In addition, in order to deepen knowledge in the field of Data Science and of Advanced Robotics, students may face: (i)topics regarding analysis and management of Big Data, visual analytics, business analytics; (Ii) or topics on computer vision, intelligent Mobile robotics, cloud robotics.<br><br>- students faces also supplementary teaching in telecommunications ( especially on virtual networks, architectures and cloud services, and wireless networks), which are important to set the educational offer in relation to specifics demands that have emerged in the context analysis.<br><br>- An important part of the training process is the use of various specialized laboratories, which allow students to deepen practical aspects through project activities, independently or by working in team.<br><br>- The student also has the possibility to conduct internships and thesis in national or international companies/universities/research centre in order to experience and extend their practical skills, and to encourage future integration into the working world. The training activities are typically concentrated in the second year, towards the end of the training course.<br><br>The final exam, usually based on a project activity, involves the preparation of a thesis about a subject in one or more areas of interest of the Board of the master degree course.<br><br>To broaden their experiences and their culture, students can access to internationalization programs, performing abroad, at universities or prestigious research centers, a part of their studies (exams or thesis),<br><br>At the end of their training, graduates of the course will be able to operate, design and produce complex and innovative systems in several major application areas, including for example: social sciences (social network analysis), IT security and national security (detection and cyber attack treatment and analysis of criminal networks), information systems and computer networks (management, processing and transmission of large data sets), computer systems (design and software development environment) |
| *Registration Deadline* | 5 November 2017 |
| *Payment* | EUR |

| Field Name | Description |
|---|---|
| *Title* | **Data Science Master Certificate** |
| *Track Name* | N/A |
| *Course List* | http://masterds.unipg.it - |
| *Organizer* | University of Perugia - Sponsorship from Engineering, IBM |
| *Type of Program* | MSc |
| *Location* | Perugia (Italy) |
| *Start Date and Time* | 2 February 2017 |
| *End Date and Time* | 5 November 2017 |
| *URL* | http://www.unipg.it/didattica/procedure-amministrative/accesso-corsi-numero-programmato/master?layout=concorso&idConcorso=838 |
| *Contact* | Gianluca Reali, reali@unipg.it |
| *Language* | English and Italian |
| *Level* | The level of studies following either Bologna or US approach |
| *Credit* | Recommended for academic courses, including grading system |
| *Prerequisites* | Master's Degree or equivalent academic title. For graduates possessing Master's Degrees in Information and Industrial Engineering (Classes from LM-20 to LM-22 and from LM25 to LM-34), Computer Sciences (LM-18), Mathematics (LM-40), Physics (LM-17) and Statistics (LM-83), |
| *Target Audience* | Students and practitioners |
| *Knowledge Areas* | DSA, DSE, DSBP |
| *Competence Groups* | DSDA01, DSDA02, DSDA04, DSDS06; DSENG03, DSENG06; DSDM02; DSDM05; DSBA01, DSBA03 |
| *Learning Outcomes* | N/A |
| *Professional Profiles* | DSP04, DSP01, DSP06, DSP08 |
| *Description* | The master, organized by the **Department of Engineering, University of Perugia,** aims to train **Data Scientist**, the new data professional. This practitioner, as a data expert, must be able to create new business strategies in statistic, computer science, economic and management fields; today, the purchase of such multidisciplinary competences are, those mainly request for the corporate data developmentand enhancement. |
| *Registration Deadline* | 10 January 2017 |
| *Payment* | 3000 Eur |

## Lucerne School of Information Technology (HSLU)

**Major in Data Engineering and Data Science**

- Describing the program and listing the courses

The first step was to describe the program and to list the courses, see chapter 2. Based on this catalogue, the learning objectives could be gathered. This involved communication with many course responsibles and a coordination within a committee (Fachgruppe Data Science & Engineering).

The Lucerne School of Information Technology offers Bachelor's degree programs in information technology and business informatics. For both programs students choose a major, that is a specialization in the third year of study in a specific IT topic of 24 ECTS. Starting in September 2016, we offer a major in *data engineering and data science*, for both BSc in information technology and in business informatics. This major requires students to choose the courses statistics and NoSQL Databases as a prerequisite. On the intermediate level (second year of study for full time students) On the major level (third year of study for full time students) the following courses are taught:

- Contact information (university, unit, contact person with email)
    - Lucerne University of Applied Science and Arts
      School of Information Technology
      Prof. Dr. Michael Kaufmann
      m.kaufmann@hslu.ch
- General information about program
- We are designing a new program
- The School of Information Technology owns the program
- Two Bachelor's degree programs with 25 ECTS Major in Data Engineering and Data Science
    - Bachelor of Science Hochschule Luzern/FHZ in Wirtschaftsinformatik, (WI)
    - Bachelor of Science Hochschule Luzern/FHZ in Informatik, Major in Data Engineering and Data Science (I)
- List of tracks, courses (other than standard I or WI courses)
- Mandatory prerequisites:
    - Statistics I
    - Statistics II
    - Database Systems (I) or Data Management + NoSQL Databases (WI)
- Major modules
    - Data Science Basics
    - Big Data Lab
    - Data Warehousing
    - Business Intelligence
    - Machine Learning
    - Text- & Social Media Analytics
- We have a project track, where at least 42 ECTS are problem-based.

## • Initial EDSF Reflection

- Desk-based, assessing scientific and technical soundness, scope and relevance, readability, appearance and structure, usability, etc.

The EDSF seems technically and theoretically very sound, as its BOK is aligned with existing books and specific data science competencies are added based on the needs of real world entities by analysis of job offerings. Also, EDSF addresses the scientific method as a guideline for data science professionals. The EDSF is highly relevant, because the need for data science education is growing, the demand in the industry is there and it is not matched by educational offerings. There is a need for structure, guidelines, frameworks, and model curricula in the field of Data Science.

The scope is okay, but rather broad in range. A more focused description for end-users could be valuable. The readability is questionable, because the mentioned documents (see prerequisites) have clearly a research character. The whole process of mental and social construction of a new academic field is visible, rather than only the condensed results of that process. Therefore, to enhance the usability of the EDSF, I suggest to provide a new additional document for end users (Universities that only want to apply the results. That document should contain only a summary of all relevant results

(i.e., without distracting descriptions of the path to these results), together with specific paths to apply thesis results.

- Expectations w.r.t. applying EDSF (Use Case description)?

We expect to use EDSF as a measure of quality management for designing our new major in data engineering and data science. By aligning our courses with an international benchmark, we can identify and close gaps and thus improve the quality of our offering. Since the new Degree programmes started in 2016, the major level will be tought from 2018. The course outline and learning objectives have been defined, but now we still have time to fine-tune the details and to add relevant objectives if there are significant gaps w.r.t. EDSF.

## • Listing the internally planned Learning Objectives and mapping to EDISON

Our program is still being set up, and first courses will start in Autumn 2018. Gathering all the learning objectives for the mapping meant asking the responsible persons of the course if they already have a draft version of the course concept including learning objectives. Also, it became clear that the program might profit even more from the application of EDSF than we thought before.

### • Statistics I

LO1.06 I

LO01.02 I, II

- F1: The students present data with the most suitable graphical representation.
  - LO1.06 I
- F2: They calculate mean values, scattering masses and index numbers with adequate tools.
  - LO01.02 I
- F3: Students know the method of the least squares. They formulate and calculate models and determine trends from time series.
  - LO01.02 I
- M1: Students create tables and graphs using a spreadsheet program such as LibreOffice or Excel.
  - LO01.02 I
- M2: You interpret the collected data correctly.
  - LO01.02 II
- P1: The students formulate statistical results for the recipient.
  - LO01.02 II
- P2: They reflect on the experience of group work.
  - -

### • Statistics II

- F1: Students know that sample-based statements are subject to a random uncertainty.
  - -
- F2: They calculate confidence intervals and perform statistical hypothesis tests.
  - LO01.02 I, LO4.01. I
- M1: Students interpret results correctly. They draw the right conclusions.
  - LO01.02 II
- P1: The students reflect on the experience gained during group work.
  - -

- Database Systems (I) or Data Management + NoSQL Databases (WI)

LO3.04 I, II, III

LO2.02 I, II

- Relational databases:
- F3: Convert a conceptual data model to a relational schema. (RS)
- F4: Normalize data models. (NF)
- F5: Understand the principles of Relational Algebra. (R1)
- F6: Apply Relational Algebra operators. (R2)
- F7: Apply SQL basics. (S1)
- F8: Apply advanced SQL. (S 2)
- F9: Use programs from databases. (DP)
- F10: Ensure semantic integrity of a database. (DI)
- F11: Ensure data security in databases. (DS)
- F12: Maintain consistency in database transactions. (TR)
- F13: Relational expressions are converted equivalently. (R3)
- F14: Optimize the performance of database queries. (AO) Postrelation databases:
- F15: Understanding multi-dimensional databases. (MD)
- F16: Apply object-relational Mapper. (OM)
- NoSQL databases:
- F17: Explain Big Data and show appropriate strategies and tools. (BD)
- F18: Define and motivate NoSQL databases. (ND)
- F19: Implement a graphical database. (GD)
- F20: Implement a document database. (DD)
- F21: Implement a key-value database. (SW)
- F22: Implement a column family database. (SF)

- Data Science Basics

LO1.01 I II III

- F1: Explain what is meant by data science today
- F2: describe the different tasks of a Data Engineer and a Data Scientist
- F3: Demonstrate different methodologies / models in the data science environment
- F4; The programming language R for problems in the DE & DS environment
- F5: Connect to a BigData platform (Hadoop-Zoo) and make simple analyses in R
- F5: Evaluate ethical concerns in Big Data Projects
- M1: Integrate autodidactically into a new technology
- M3: Organizing a project and meeting deadlines meet the goals
- M4: summarize and report on project results
- P1: Organize yourself in the group
- P2: Hold a presentation in front of an audience
- P3: submit work results on time
- P4: Give peer feedback
- P5: Accept peer feedback

- Big Data Lab

LO1.04 I, II, III

LO1.05 I, II, III

LO3.06 III

- F1: Get to know and use proven and up-to-date tools in the area of BigData and NoSQL.
- F2: You can assess which tools are best suited for which questions.
- F3: Deepen in at least one area independently and present this knowledge.
- F4: Deepening into a new little-known tool in the BigData / NoSQL environment and presenting this knowledge.

- Data Warehousing

LO2.02 I, II, II

L02.03 I

LO2.04 I, II, III

- F1: The students know the basic concepts, models and architectures of analytical data stocks.
- F2: Students are able to model analytical data (from the operational side).
- F3: Students can choose the models that are most problem-oriented
- F4: Students know the problems of denormalization and data quality in heterogeneous data collection.
- F5: Students analyze structured and unstructured data.
- F6: Students know the metadata modeling.
- M1: Students know the big picture of data warehousing.
- M2: The students know the most important tools for (transformation into) analytical data sets.
- M3: Students use the right DWH tool in the right place at the right time.
- P1: The students know and recognize data warehousing not as a self-purpose but as an instrumental discipline in a business environment.
- P2: Students are aware of the problem of collecting and evaluating personal data.
- P3: The students act as competent, socially conscious interlocutors and co-decision-makers in decisions about data warehousing.

- Business Intelligence

LO5.01 I, II, III

LO01.03 I, II

- F1: Students know the basics and processes of business intelligence.
- F2: Students are able to determine information needs and to implement appropriate processes for obtaining the information.
- F3: The students know the typical fields of application of Business Intelligence.
- F4: The students can use common procedures in the treated areas.
- M1: The students know the most important methods of business intelligence.
- M2: Students use the appropriate means of Business Intelligence to develop adequate decision-making tools.
- M3: The students know which problems and solutions are used in the typical fields of application.
- M4: The students are able to use the methods described adequately.
- P1: Students are familiar with the way in which information is designed to be user-friendly. They can convey knowledge gained in a suitable way.

- P2: Students know that informed decisions depend on adequate information and are aware of the impact of incorrect or inadequate information on decision-making.
- P3: Students understand the hurdle of BI results and the communication of the findings in everyday life. They are able to com- municate the acquired knowledge competently.
- P4: Students will see how the procedures in the treated areas can be transferred to other areas. They make proposals for the use of BI techniques in adjacent fields of application.

### • Machine Learning

LO1.03 I, II, III

- Describe a typical task of data analysis as a machine learning problem
- Suggest an adequate machine learning procedure for a given problem situation and implement it using a framework
- From the result of a machine learning procedure, draw correct conclusions on the original problem situation
- Plan and evaluate the evaluation of a machine learning process and present the derived findings correctly and in detail
- Assess the quality of a machine learning process, assess and, if necessary, propose improvements

### • Text- & Social Media Analytics

LO3.06 II

LO1.04 II

- F1: explaining basic concepts and connections in the field of text and social media analytics
- F2: Unstructure and use unstructured text from the web, social media, and document collections using the appropriate tools (Twitter API, Nutch, Solr, Tika, Lucene, MoreLikeThis)
- F3: Access social media data via Twitter API
- F4: Parsing a text source (the web, a document base) with Apache Nutch
- F5: Extract full text from Office, PDF, HTML, etc. with Apache Tika
- F6: Indexing in Apache Lucene and retrieving it via search queries
- F7: With MoreLikeThis, the texts indexed in Lucene are automatically tagged with key words and related documents
- F8: Programming a simple text categorization and recommender system that extracts relevant keywords based on a document and suggests related, similar documents
- M2: Integrate autodidactically into a new technology
- M3: Organizing a project and meeting deadlines meet the goals
- M4: summarize and report on project results
- P1: Organize yourself in the group
- P2: Hold a presentation in front of an audience
- P3: submit work results
- P4: Give peer feedback
- P5: Accept peer feedback

### • Data Visualization

L01.06 I. II. III

- F1: Media Configurations: Which (visual) media are suitable for which form of communication?

- F2: perception and cognition: visual variables ( "preattentive variables"), gestalt laws, affordances, mental models, scripts, storyboards
- F3: Representation context: explanatory -> presentation; Exploratory -> interactive analysis; Rhetorical -> conviction
- F4: IBCS communication standards: Knowledge and effective application of international business communication standards
- F5: Tufte's Guidelines: 1. Graphical Excellence; 2. Visual Integrity; 3. Maximizing the data-ink ratio; 4. Aesthetic Elegance
- F6: Interactive data visualization: Knowledge and practical implementation of the Shneiderman mantra: overview first, filter, detail on demand
- F7: Representation forms for data structures: hierarchical, relational (network), temporal, spatial (geospatial), spatiotemporal, textual
- F8: Introduction Cartography: Advantages and disadvantages of different projection possibilities, metadata, layer concepts.
- F9: Dashboard design
- F10: Historical retrospect

## Gathering the relevant Information of the EDSF

For the matching and gap analysis, the relevant learning objectives, competence groups, etc, from the edsf had to be gathered. First, the Documents mentioned in Section 1 were downloaded. Because the terminology of EDISON is rather cryptic, the actual deliverables were classified according to the following structure, and relevant corresponding EDSF content was assigned to get an overview.

Because the deliverables were not provided in word format, extracting relevant information was not straight-forward. The necessary pages were cut out with Acrobat Pro and exported to Word Format. Thus, a summarized handbook was created that facilitated the following mapping. This file is available as an annex to this document.

## Mapping to the EDISON resources

The following questions have to be answered by the case study:

### CF-DS: Specification of matching competence groups, competences... and learning outcomes

The following competence groups, competences and learning outcomes match: (details, see the list of our courses, learning objectives, and mapping to EDISON above. The mapping has been encoded there on the detailed level of learning outcomes)

Legend:

Learning outcome level I level II level III

| | | | | |
|---|---|---|---|---|
| DSDA01<br>Use predictive analytics to analyse big data and discover new relations<br>Data Science Basics<br>Machine Learning | DSDM01<br>Develop and Implement data strategy, in particular, in a form of Data Management Plan (DMP) | DSENG01<br>Use engineering principles to research, design, prototype, data analytics applications, or develop structures, instruments, machines, experiments, | DSRM01<br>Create new understandings and capabilities by using the scientificmethod (hypothesis, test, and evaluation) or similar engineering research and development methods | DSDK01<br>Understand business and provide insight, translate unstructured bu siness problems into an abstract mathematical framework<br>Business Intelligence |
| DSDA02<br>Use appropriate statistical techniques on available data to deliver insights<br>Satistics I / II | DSDM02<br>Develop and implement relevant data models, including metadata<br>Database Systems<br>Data Warehousing | DSENG02<br>Develop and apply computational solutions to domain related problems using wide range of data analytics platforms<br>Big Data Lab | DSRM02<br>Direct systematic study toward a fuller knowledge or understanding of the observable facts, and discovers new approaches to achieve research or organisational goals | DSDK02<br>Use data to improve existing services or develop new services |
| DSDA03<br>Develop specialized analytics to Enable agile decision making<br>Business Intelligence<br>Machine Learning | DSDM03<br>Collect and integrate different data source and provide them for further analysis<br>Data Warehousing | DSENG03<br>Develops specialized data analysis tools to support executive decision making | DSRM03<br>Undertakes creative work, making systematic use of investigation, experimentation, discover or revise knowledge of reality, and uses this knowledge to devise | DSDK03<br>Participate strategically and tactically in financial decisions that impact management and organizations |
| DSDA04<br>Research and analyse complex datsets, combine different sources andtypes of data to improve analysis.<br>Big Data Lab | DSDM04<br>Develop and maintain a historical data repository of analysis results (data provenance)<br>Data Warehosing | DSENG04<br>Design, build, operate relational non-relational databases<br><br>Database Systems | DSRM04<br>Ability to translate strategies into action plans and follow through to completion. | DSDK04<br>Provides scientific, technical, and analytic support services to other organisational roles |
| DSDA05<br>Use different data analytics platforms to process complex data<br>Big Data Lab | DSDM05<br>Ensure data quality, accessibility, publications (data curation) | DSENG05<br>Develop solutions for secure and reliable data access | DSRM05<br>Contribute to and influence the development of organizational objectives | DSDK05<br>Analyse customer data to identify/optimise customer relations actions |
| DSDA06<br>Visualise complex and variable data.<br>Statistics I<br>Data Visualization | DSDM06<br>Manage IPR and ethical issues in data management | DSENG06<br>Prototype new data analytics applications<br>Text & Social Media Analytics<br>Big Data Lab | DERM06<br>Apply ingenuity to complex problems, develop innovative ideas | DSDK06<br>Analyse multiple data sources for marketing purposes |

- **Balance of program w.r.t. Competence groups (what percentage of ECTS points per group)?**
  - DSDA 44%

- DSDM 21%
- DSDE 23%
- DSRM 3%
- DSDK 9%

- **Are learning outcomes defined? Do they follow Bloom's Taxonomy?**

Learning outcomes are clearly defined. They do not follow Bloom's Taxonomy. We classfy our learning outcomes into the groups professional competence, method competence and person competence.

- **If learning outcomes defined, how do they align with EDISON learning outcomes in content and relative amount of ECTS points.**

See table above

- **Issues encountered and feedback**

The excel is too cumbersome. It is already quite an abstract task to generate a mapping. I think free text should suffice.

The numbering of all items is good, it can be used as reference

It is good that competence groups and knowledge area groups correspond. However, the order is different, which makes it confusing.

Also, there are redundant knowledge units.

It was difficult to gather all the relevant information in the different EDISON docs. It would be helpful to summarize the main points in a short Handbook. I send you an example.

- BoK-DS: Specification of matching knowledge areas, topics

The following knowledge areas match (marked in green)

- KAG1-DSDA: Data Analytics

| **Knowledge Areas (KA)** | **Suggested Knowledge Units (KU)** |
|---|---|
| Theory of computation | Design and Analysis of Algorithms |
| | Machine Learning Theory |
| | Game Theory & Mechanism design |
| | EXTENSIBILITY Point: Theory of computation |
| Mathematics of computing | Discrete Mathematics and Graph Theory |
| | Probability & Statistics |
| | Probabilistic reasoning |
| | Statistical methods, including descriptive statistics, exploratory data analysis (EDA) and confirmatory data analysis (CDA) |
| | Information theory |
| | Mathematical analysis |
| | Mathematical software and tools |

| | |
|---|---|
| | EXTENSIBILITY Point: Mathematics of Data Science (computing) |
| Computing methodologies | Artificial Intelligence |
| | Natural Language Processing |
| | Knowledge Representation and Reasoning |
| | Data mining and knowledge discovery |
| | Text analysis, Data mining |
| | Text analytics including statistical, linguistic, and structural techniques to analyse structured and unstructured data |
| | Machine Learning theory and algorithms |
| | Classification methods |
| | EXTENSIBILITY Point: Computing methodologies |
| Information systems (to support Data Science applications) | Decision Analysis and Decision Support Systems |
| | Data warehousing and Data Mining |
| | Data Analysis and statistics |
| | Multimedia information systems |
| | Data Mining |
| | Predictive analytics and predictive forecasting |
| | EXTENSIBILITY Point: Information systems |
| Big Data Technologies and Systems | Big Data algorithm for large scale data processing |
| | Big Data Analytics |
| | Big Data systems |
| | Big Data algorithms for data ingest, pre-processing, and visualisation |
| | Big Data analytics platforms and tools (including Hadoop, Spark, and cloud based Big Data services) |
| | Big Data systems for application domains |
| | EXTENSIBILITY Point: Information systems |
| Computer systems organisation for Big Data applications (including high performance networks) | Parallel and Distributed Computer Architecture |
| | Computer networks: architectures and protocols |
| | Computer networks for high- performance computing and Big Data infrastructure |
| | EXTENSIBILITY Point: |
| Big Data software organisation and engineering | Software (systems) architectures |
| | Requirements engineering and software systems development |
| | Large and ultra-large scale software systems organisation |
| | Cloud enabled applications development |

- KAG2-DSENG: Data Science Engineering

| **Knowledge Areas (KA)** | **Suggested Knowledge Units (KU)** |
|---|---|
| Computer systems organisation for Big Data applications (including high performance networks) | Parallel and Distributed Computer Architecture |
| | Computer networks: architectures and protocols |
| | Computer networks for high- performance computing and Big Data infrastructure |
| | EXTENSIBILITY Point: |
| Big Data software organisation and engineering | Software (systems) architectures |
| | Requirements engineering and software systems development |
| | Large and ultra-large scale software systems organisation |
| | Cloud enabled applications development |
| | |
| | EXTENSIBILITY Point: |
| Modelling and simulation | Modelling and simulation theory and techniques (general and domain oriented) |
| | Large scale modelling and simulation systems |
| | EXTENSIBILITY Point: |
| Big Data systems organisation and management | Enterprise information systems |
| | Large scale data storage and data management systems |
| | Collaborative and social computing systems and tools |
| | EXTENSIBILITY Point: |
| Big Data (Data Science) applications design | Programming languages for Big Data analytics: R, python, others |
| | Models and languages for complex interlinked data presentation and visualisation |
| | EXTENSIBILITY Point: |
| Infrastructure and platforms for Data Science applications group: | Cloud Computing architecture and services |
| | Cloud Computing Engineering (infrastructure and services design, management and operation) |
| | Big Data and cloud based systems design and development |
| | Cloud based applications and services operation and management |
| | Big Data Analytics platforms (including cloud based) |

| | |
|---|---|
| | Big Data Infrastructure: services and components, including data storage infrastructure |
| | Data security and protection |
| | EXTENSIBILITY Point: |
| Software engineering and management | Software requirements and design |
| | Software engineering models and methods |
| | Software quality assurance |
| | Agile development methods, platforms and tools |
| | DevOps and continuous deployment and improvement paradigm |
| | EXTENSIBILITY Point: |

- KAG3-DSDM: Data Management

| **Knowledge Areas (KA)** | **Suggested Knowledge Units (KU)** |
|---|---|
| Data management systems | Database management systems |
| | Database design and models |
| | Data Modelling, Databases and Database Management Systems |
| | Data Models and Query Languages |
| | Database administration |
| | EXTENSIBILITY Point: |
| Digital libraries and archives | Digital libraries and archives organisation |
| | Information Retrieval |
| | Data curation and provenance |
| | Search Engines technologies |
| | EXTENSIBILITY Point: |
| Data Management and Enterprise data infrastructure | Data management, including Reference and Master Data |
| | Data Warehousing and Business Intelligence |
| | Data storage and operations |
| | Data archives/storage compliance and certification |
| | Metadata, linked data, provenance |
| | Data infrastructure, data registries and data factories |
| | Data security and protection |
| | Data governance, data quality, data Integration and Interoperability |

| | |
|---|---|
| | Data Management Planning |
| | Responsible data use, data privacy, ethical principles, legal issues |
| | EXTENSIBILITY Point: |
| General principles and concepts in | Data type registries, PID, metadata |
| Data Management and organisation | Research data infrastructure, Open Science, Open Data, Open Access, ORCID |
| | Data infrastructure compliance and certification |
| | Ethical principle and data privacy |
| | EXTENSIBILITY Point: |

- KAG4-DSRM: Scientific and Research Methods

| **Knowledge Areas (KA)** | **Suggested Knowledge Units (KU)** |
|---|---|
| Scientific/Research Methods | Research methodology, paradigms and research cycle |
| | Modelling and experiment planning |
| | Data selection and quality evaluation |
| | Use cases analysis: research infrastructures and projects |
| | Research data management plan and ethical issues |
| | EXTENSIBILITY Point: |

- KAG5-DSBPM: Business process management

| **Knowledge Areas (KA)** | **Suggested Knowledge Units (KU)** |
|---|---|
| Business Process Management | Business processes and operations |
| | Project scope and risk management |
| | EXTENSIBILITY Point: |
| Business Analysis organisation and management | Business Analysis Planning and Monitoring |
| | Requirements Analysis and Design Definition |
| | Requirements Life Cycle Management (from inception to retirement) |
| | Solution Evaluation and improvements recommendation |
| | EXTENSIBILITY Point: |
| Business analysis and enterprise organisation | Agile Data Driven methodologies, processes and enterprises |
| | Use cases analysis: business and industry |
| | EXTENSIBILITY Point: |

## • Assignment of professional profiles targeted by program

The following professional profiles are targeted by our programs: DSP03, DSP06, DSP07, DSP14, DSP15, DSP17

## • MC-DS: Curricula Design and/or Evaluation

- **Gap analysis (remember to consider professional profiles, for existing programs or for using existing courses for new programs).**

See table above

- **What are the gaps on the level of balance between competence groups?**

Our program is for BSC IT and Business IT. We focus on Data Science Analytics and Engineering. The other competence groups, especially research method and domain knowledge, we have large gaps

- **What are the gaps on the level of learning outcomes?**

See table above. Learning outcomes and competencies match, the only difference is the expertise level, and this is encoded using the colors in the table,

# • Our Approach

The simple approach was to to align the process of application of the EDSF with our own eixisting curriculum structure. We took the following steps:

- Describing our program.
- Describing the initial reflection
- Gathering the relvant information on our internal courses including learning outcomes
    - This was time consuming because we had to coordinate with all professors and lecturers that teach in the program and also with the directors
- Gathering the relevant information on EDISON.
    - This was abstract and complex, because the information is hidden within four different documents. We had to search and find the relevant pieces. Copying it into a summary document was difficult because the files were in PDF, not Word
- Map the EDISON resources with our program.
    - Every single course was mapped to competencies and learning outcomes. Competencies and groups can be directly derived from the learning outcomes. There is a mapping and corresponcencebetween them.
- Gap analysis.
    - Instead of using the Excel, which are cumbersome and inefficient, we used the overview table for competencies and groups CFDS Table 4.1 and color-coded it to summarize the mapping result. This is also much more human readable, as the gaps can be seen easily.
- Description of our approach
    - The approach is also reflected in the document structure.

This task turned out to be quite

# • Summary and Outlook

We will discuss the results in the Fachgruppe Data Science & Engineering. We will discuss to include more compentencies of EDISON in our Curriculum.

Since our program is a specialization of existing BSC IT and business IT, we cannot cover all the topics. However, we are currently evaluating to implement a Master's program for applied data science. There we use the EDISON guidelines for curriculum development. The program is not ready to be matched based on the level of individual learning outcomes. Yet, we have structured our initial concept starting with EDISON professional profiles mapped to competency groups, and we have identified possible courses using the EDISON Knowledge units. The framework helps us greatly to structure the formidable task of designing a new program from scratch. Aligning it with an existing EU framework gives our concept more credibility, it simplifies the task so that we can choose from a catalogue of elements, and this ensures the quality because we apply excellent preliminary work.

| **Field Name** | **Description** |
|---|---|
| *Title* | Bachelor of Science Hochschule Luzern/FHZ in Wirtschaftsinformatik |
| ***Track Name*** | **Major in Data Engineering and Data Science** |
| *Course List* | https://www.hslu.ch/de-ch/informatik/studium/bachelor/majors/kernmodule/%20#kernmoduledataengineering |
| *Organizer* | Lucerne University of Applied Sciences and Arts, School of Information Technology |
| *Type of Program* | Academic program |
| *Location* | Rotkreuz, Switzerland |
| *Start Date and Time* | Fall 2016 |
| *End Date and Time* | N/A |
| *URL* | N/A |
| *Contact* | m.kaufmann@hslu.ch |
| *Language* | German |
| *Level* | Bachelor |
| *Credit* | 180 ECTS |
| *Prerequisites* | Matura (A-Level) plus 1 year of professional experience in IT |
| *Target Audience* | IT professionals |
| *Knowledge Areas* | Described above |
| *Competence Groups* | Described above |
| *Learning Outcomes* | Described above |
| *Professional Profiles* | Described above |
| *Description* | Described above |
| *Registration Deadline* | N/A |
| *Payment* | N/A |

The courses Statistics I, Statistics II, Database Systems, Data Science Basics, Big Data Lab, Data Warehousing, Machine Learning, Business Intelligence are described in detail in the walkthrough above.
They will all have a structure like the following:

| **Field Name** | **Description** |
|---|---|

| *Title* | **Statistics I** |
|---|---|
| *Name of Presenter(s)* | N/A |
| *Organizer* | Lucerne University of Applied Sciences and Arts, School of Information Technology |
| *Type of Course* | Academic Course |
| *Related Program* | Recommended |
| *Location* | Rotkreuz, Switzerland |
| *Start Date and Time* | Mandatory |
| *End Date and Time* | Mandatory |
| *URL* | Mandatory |
| *Contact* | m.kaufmann@hslu.ch |
| *Language* | German |
| *Level* | Bachelor's Degree |
| *Credit* | 3 ECTS |
| *Prerequisites* | N/A |
| *Target Audience* | Optional |
| *Knowledge Areas* | Described above |
| *Competence Groups* | Described above |
| *Learning Outcomes* | Described above |
| *Description* | Described above |
| *Registration Deadline* | N/A |
| *Payment* | N/A |

**Master of Arts in Applied Data Science**

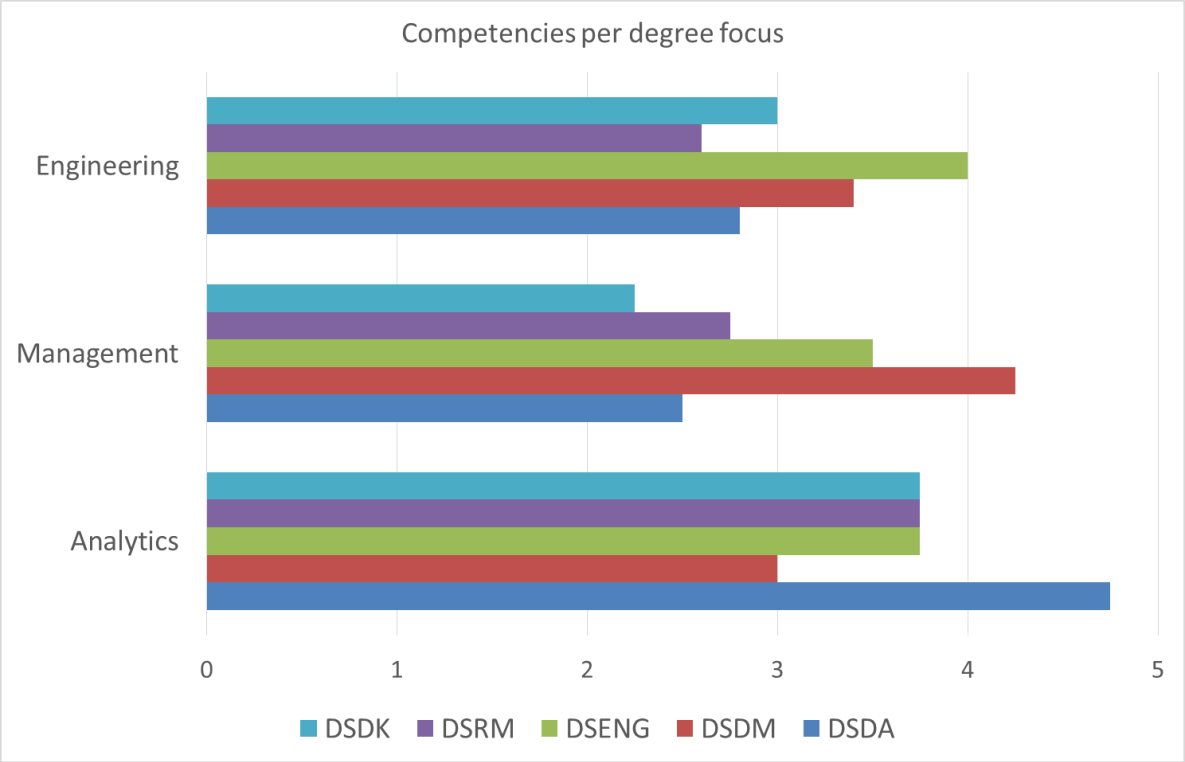
Additional Case Study of Application of the EDSF on a newly developed Master Program Applied Data Science. The development is still ongoing.

**Mapping Competencies:**

| | | | | | | | Master of Arts in Applied Data Science | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | |
| **Managers** | DS DA | DS DM | DS ENG | DS RM | DS DK | | Analytics | Management | Engineering |
| DSP01 Data Science (group) Manager | 3 | 4 | 3 | 3 | 2 | | | x | |
| DSP02 Data Science Infrastructure Manager | 2 | 4 | 4 | 2 | 2 | | | x | |
| DSP03 Research Infrastructure Manager | 2 | 4 | 4 | 3 | 2 | | | x | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Professionals** | | | | | | | | | |
| DSP04 Data Scientist | 5 | 3 | 4 | 5 | 3 | | x | | |
| DSP05 Data Science Researcher | 4 | 3 | 2 | 5 | 4 | | | | |
| DSP06 Data Science Architect | 4 | 3 | 5 | 3 | 3 | | x | | x |
| DSP07 Data Science (Application) Programmer/Engineer | 4 | 2 | 5 | 3 | 3 | | | | x |
| DSP08 Data Analyst | 5 | 3 | 3 | 3 | 4 | | x | | |
| DSP09 Business Analyst | 5 | 3 | 3 | 4 | 5 | | x | | |
| **Professional (data handling/ management)** | | | | | | | | | |
| DSP10 Data Stewards | 3 | 5 | 3 | 3 | 3 | | | x | |
| DSP11 Digital data curator | 1 | 5 | 2 | 2 | 3 | | | | |
| DSP12 Digital Librarians | 2 | 5 | 2 | 2 | 3 | | | | |
| DSP13 Data Archivists | 1 | 5 | 1 | 1 | 3 | | | | |
| **Professional (database)** | | | | | | | | | |
| DSP14 Large scale (cloud) database designer | 2 | 4 | 4 | 3 | 3 | | | | x |
| DSP15 Large scale (cloud) database administrator | 2 | 4 | 3 | 2 | 3 | | | | x |
| DSP16 Scientific database administrator | 2 | 4 | 3 | 2 | 3 | | | | x |
| **Technicians and associate professionals** | | | | | | | | | |
| DSP17 Big Data facilities Operator | 1 | 4 | 3 | 2 | 3 | | | | |
| DSP18 Large scale (cloud) data storage operator | 1 | 4 | 3 | 1 | 1 | | | | |
| DSP19 scientific database operator | 1 | 4 | 3 | 2 | 3 | | | | |
| **Clerical support workers (general and keyboard workers)** | | | | | | | | | |
| DSP20 data entry/access workers | 2 | 1 | 0 | 0 | 2 | | | | |
| DSP21 data entry field workers | 2 | 1 | 0 | 0 | 2 | | | | |
| DSP22 User support data services | 3 | 2 | 0 | 0 | 2 | | | | |

| | DSDA | DSDM | DSENG | DSRM | DSDK | |
|---|---|---|---|---|---|---|
| Analytics | 4,75 | 3 | 3,75 | 3,75 | 3,75 | 0 |
| Management | 2,5 | 4,25 | 3,5 | 2,75 | 2,25 | |
| Engineering | 2,8 | 3,4 | 4 | 2,6 | 3 | |




Data Science Analytics (DSDA), Data Management (DSDM), Data Science Engineering (DSENG), Scientific/ Research Methods (DSRM), DS Domain Knowledge, e.g., Business Apps (DSDK)

**Mapping Knowledge Units:**

| CG / KAG | Target competence | Title | Knowledge areas (possible modules) |
|---|---|---|---|
| KAG1-DSDA | DSDA01 | Predictive Analytics | Data mining<br>Text Analytics<br>Information Retrieval |
| KAG1-DSDA | DSDA02 | Statistics | Statistical Methods<br>Data Query & Manipulation Languages |
| | DSDA04 | Complex Data Analysis | Big Data Analytics<br>Artificial Intelligence<br>Machine Learning<br>Natural Language Processing |
| KAG1-DSDA | DSDA06 | Visualization | Data Visualization<br>Data Communication |
| KAG2-DSENG | DSENG01 | Engineering Principles | Data Engineering |
| KAG2-DSENG | DSENG02 | Computational Thinking | Computational Thinking for Data Analytics |

| | | | |
|---|---|---|---|
| KAG2-DSENG | DSENG03 | DSS Development | Decision Support Systems<br>Business Intelligence<br>Management Support Systems<br>Enterprise Information Systems<br>Search Engine Technology |
| KAG2-DSENG | DSENG04 | DBMS | Relational Databases<br>NoSQL Databases<br>Large Scale data systems |
| KAG2-DSENG | DSENG05 | Data Security | Data security management<br>Data Protection |
| KAG2-DSENG | DSENG06 | DS Application Prototyping | Programmig Languages for Data Analytics<br>Cloud Computing architecture and services |
| KAG3-DSDM | DSDM01 | Data Strategy & Planning | Reference and Master Data Management |
| KAG3-DSDM | DSDM02 | Data Modelling | Data Models |
| KAG3-DSDM | DSDM03 | Data Integration / ETL | Enterprise Application Integration<br>Metadata management |
| KAG3-DSDM | DSDM04 | Data Historization | Data Warehousing |
| KAG3-DSDM | DSDM05 | Data Quality | Data Quality Management |
| KAG3-DSDM | DSDM06 | Legal / ethical | Data privacy<br>Ethics of Data Science |
| KAG4-DSRM | DSRM01 | Scientific Method | Scientific Method for Data Analysis |
| KAG4-DSRM | DSRM01 | Design Thinking | Design Thinking for Data Analytics |
| KAG5-DSRM | DSDK02 | Data based Innovation | Data based Innovation |
| KAG5-DSRM | DSDK05 | Analytic CRM | Analytic CRM |
| KAG5-DSRM | DSDK06 | Marketing Analytics | Marketing Analytics |

**Mapping Learning Outcomes:**

| **Target competences** | **Title** | **Prio** | **Expertise** |
|---|---|---|---|
| DSDA01 | Predictive Analytics | 1 | Assessment |
| DSDA02 | Statistics | 1 | Useage |
| DSDA04 | Complex Data Analysis | 2 | Familiarity |
| DSDA06 | Visualization | 1 | Useage |
| DSENG01 | Engineering Principles | 2 | Familiarity |
| DSENG02 | Computational Thinking | 1 | Useage |
| DSENG03 | DSS Development | 2 | Familiarity |

| | | | |
|---|---|---|---|
| DSENG04 | DBMS | 1 | Assessment |
| DSENG05 | Data Security | 2 | Familiarity |
| DSENG06 | DS Application Prototyping | 1 | Useage |
| DSDM01 | Data Strategy & Planning | 2 | Familiarity |
| DSDM02 | Data Modelling | 1 | Useage |
| DSDM03 | Data Integration / ETL | 1 | Useage |
| DSDM04 | Data Historization | 2 | Familiarity |
| DSDM05 | Data Quality | 2 | Familiarity |
| DSDM06 | Legal / ethical | 2 | Familiarity |
| DSRM01 | Scientific Method | 1 | Useage |
| DSRM01 | Design Thinking | 1 | Useage |
| DSDK02 | Data based Innovation | 1 | Useage |
| DSDK05 | Analytic CRM | 1 | Assessment |
| DSDK06 | Marketing Analytics | 2 | Familiarity |

## The School of Data Science

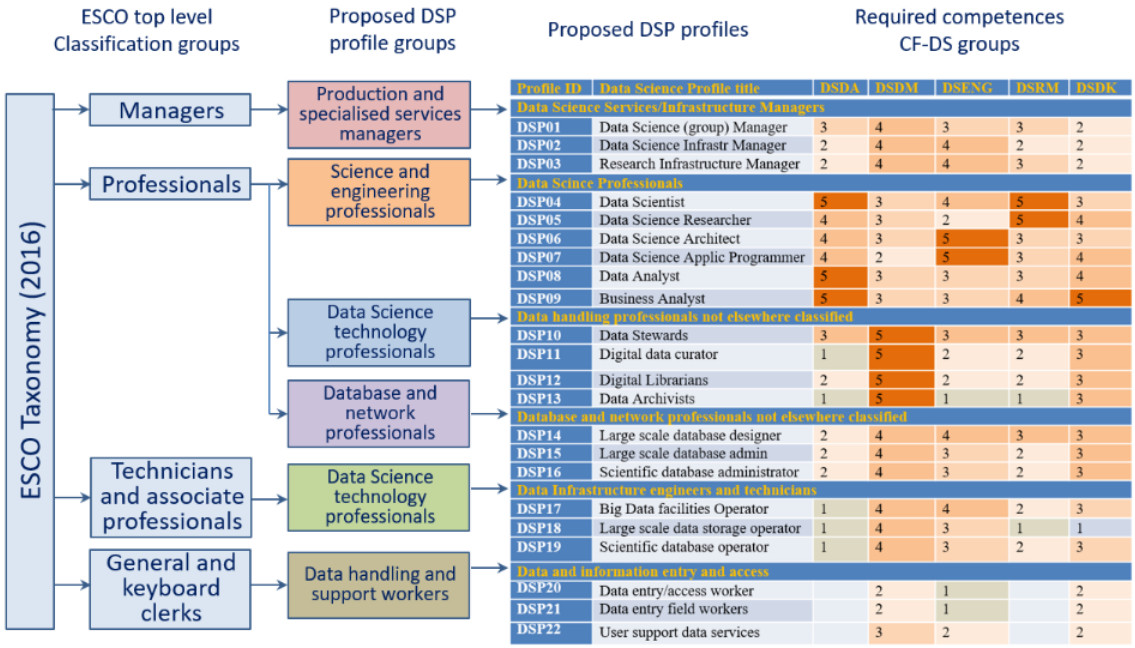
Persontyle launched their School of Data Science in 2014. The goal is to produce data science training and education for professionals. What is offered:

- Data Science/Machine Learning Courses
- Hadoop Training
- On-site Corporate Training
- 6 week Bootcamp for PhD graduates
- Fellows program for using data science for social good

The offerings are available for a reasonable fee and they are taught in cities around the globe. The training is individualized and can be adjusted to the special needs of the participants. Plus, it is targeted at businesses and working professionals.

Currently Persontyle is restructuring their offerings. The EDSF guidelines will help to especially match the professional profiles of future Data Scientists with the line-up of Persontyle.

The overview of DSPP is very useful as guideline.




| Profile ID | Data Science Profile title | DSDA | DSDM | DSENG | DSRM | DSDK |
|---|---|---|---|---|---|---|
| **Data Science Services/Infrastructure Managers** | | | | | | |
| DSP01 | Data Science (group) Manager | 3 | 4 | 3 | 3 | 2 |
| DSP02 | Data Science Infrastr Manager | 2 | 4 | 4 | 2 | 2 |
| DSP03 | Research Infrastructure Manager | 2 | 4 | 4 | 3 | 2 |
| **Data Scince Professionals** | | | | | | |
| DSP04 | Data Scientist | 5 | 3 | 4 | 5 | 3 |
| DSP05 | Data Science Researcher | 4 | 3 | 2 | 5 | 4 |
| DSP06 | Data Science Architect | 4 | 3 | 5 | 3 | 3 |
| DSP07 | Data Science Applic Programmer | 4 | 2 | 5 | 3 | 4 |
| DSP08 | Data Analyst | 5 | 3 | 3 | 3 | 4 |
| DSP09 | Business Analyst | 5 | 3 | 3 | 4 | 5 |
| **Data handling professionals not elsewhere classified** | | | | | | |
| DSP10 | Data Stewards | 3 | 5 | 3 | 3 | 3 |
| DSP11 | Digital data curator | 1 | 5 | 2 | 2 | 3 |
| DSP12 | Digital Librarians | 2 | 5 | 2 | 2 | 3 |
| DSP13 | Data Archivists | 1 | 5 | 1 | 1 | 3 |
| **Database and network professionals not elsewhere classified** | | | | | | |
| DSP14 | Large scale database designer | 2 | 4 | 4 | 3 | 3 |
| DSP15 | Large scale database admin | 2 | 4 | 3 | 2 | 3 |
| DSP16 | Scientific database administrator | 2 | 4 | 3 | 2 | 3 |
| **Data Infrastructure engineers and technicians** | | | | | | |
| DSP17 | Big Data facilities Operator | 1 | 4 | 4 | 2 | 3 |
| DSP18 | Large scale data storage operator | 1 | 4 | 3 | 1 | 1 |
| DSP19 | Scientific database operator | 1 | 4 | 3 | 2 | 3 |
| **Data and information entry and access** | | | | | | |
| DSP20 | Data entry/access worker | | 2 | 1 | | 2 |
| DSP21 | Data entry field workers | | 2 | 1 | | 2 |
| DSP22 | User support data services | | 3 | 2 | | 2 |

**Figure 12: DSPP by classification groups**

In addition the developments on team composition are great to get a new view on possible team compilations and extended fields of working for Data Scientists.
The current focus of training and education courses of The School of Data Science is primary covering Data Science and Machine Learning, so the DSENG part. The Data Management related profiles and competences would be a very interesting supplementing and will be discussed for the new offerings in 2018.