

# Current State Of Demand Of Data Science Skills & Competences

*Thursday, 12 November 2015 09:10 (20 minutes)*

The emergence of Data Science technologies (also referred to as Data Intensive Science or Big Data technologies) is having an impact, at a fundamental level, on nearly every aspect of how research is conducted, how research data are used and shared. Data Science is considered as main enabler and facilitator of the recently launched by EC the Open Science initiative for European Research Area (ERA).

The effective use of Data Science technologies requires new skills and demands for new professions, usually referred as the Data Scientist: an expert who is capable both to extract meaningful value from the data collected and also manage the whole lifecycle of Data, including supporting Scientific Data e-Infrastructures. The future Data Scientists must possess knowledge (and obtain competencies and skills) in data mining and analytics, information visualisation and communication, as well as in statistics, engineering and computer science, and acquire experiences in the specific research or industry domain of their future work and specialisation.

The Horizon 2020 EDISON project (1 September 2015 –30 August 2017) aims to develop a sustainable business model that will ensure a significant increase in the number and quality of data scientists graduating from universities and being trained by other professional education and training institutions in Europe. This will be accomplished through the development of a number of inter-connected activities including the definition of required skills and competences, defining a Data Science Competence Framework (CF-DS) and also a Model Curriculum (MC-DS).

The project will work in close co-operation with experts and practitioners involved and interested in the development of Data Science academic educational and professional training programmes. The target is to define basic competences and skills for new profession of data scientist with the focus on European e-Infrastructure and industry needs. This cooperation will take place in consultation and validation activities and roundtables with relevant stakeholders and institutions.

The proposed workshop Demand Of Data Science Skills & Competences will contribute by presented EDISON recent development and initiate an open dialogue across communities to characterize the existing needs and practical requirements, topical trends and preferences in order to create the Competence Framework for Data Science.

Participants are invited to contribute to the following objectives:

- Analysis of organizational and employer requirements for competences and skills of Data Scientists
- Elicitation of education and training needs in various contexts, focusing on different industry sectors. Particular training and continuous education needs for practitioners (or self-made data scientists) to support advanced European Research e-Infrastructure development.
- Discussing perspectives of the educational path for the Universities with the potentials of starting careers into data-driven industries, as well as for life-long learning programs to identify suitable paths to career development
- Identified profiles corresponding to the major stakeholders/employers of the future specialists

Target audience:

European Research e-Infrastructure and European research institutions, data intensive and data driven industries, innovation companies and SME as well as related community initiatives interested in supporting the developing in data science experts

## Links, references, publications, etc.

European eCompetences Framework <http://www.ecompetences.eu/>

CEN ICT skills Workshop <http://www.ecompetences.eu/cen-ict-skills-workshop/>

Demchenko, Yuri, David Bernstein, A.S.Z. Belloum, Cees de Laat, Experience of Profiling Curricula on Cloud Computing Technologies and Engineering for Different Target Groups , IEEE International Conference on Cloud Computing Technology and Science, (CloudCom 2014), Singapore, Dec. 2014.

Demchenko, Yuri, David Bernstein, A.S.Z. Belloum, Ana Oprescu, Tomasz W. Włodarczyk, and Cees de Laat, New Instructional Models for Building Effective Curricula on Cloud, Computing Technologies and Engineering,

IEEE International Conference on Cloud Computing Technology and Science, Bristol, Dec 2013.

Skills and Human Resources for e-Infrastructures within Horizon 2020, The Report on the Consultation Workshop, May 2012

## **Additional information**

The EDISON project as initiator of this workshop will present ongoing work on the definition of mentioned above components CF-DS and MC-DS and will intend to build cooperative links with EGI community. The EDISON Framework needs to consider specifics of each scientific domain and research community, as well as for industry and public sector. This will be achieved by analyzing the requirements of the communities in order to identify the competencies and skills required by the Data Scientist Professionals, surveying the existing education and training resources to create an inventory and the Data Science Profession taxonomy and the profiles for the Data Scientists. All these will constitute the theoretical basis of the EDISON Framework that will be the basis and be complemented by the definition of the Model Curriculum. These activities and results are a cornerstone for the EDISON project and all other activities will depend from the high quality of the results obtained.

The workshop will also present and link to the related activity at Research Data Alliance (RDA) in particular conducted in the framework of Interest Group on Education and Training on Research Data Handling (IG-ETRD). Related representatives will be invited to contribute to workshop.

**Primary authors:** BROCKS, Holger (InConTec GmbH); BECKER, Jana (FTK - Forschungsinstitut fuer Telekommunikation und Kooperation e.V.); Prof. HEMMJE, Matthias (FernUniversität in Hagen)

**Co-authors:** MANIERI, Andrea (Engineering); DESKA-DRÜKE, Beate (FTK - Forschungsinstitut fuer Telekommunikation und Kooperation e.V.); RIESTRA, Ruben (INMARK); Dr FERRARI, Tiziana (EGI.eu)

**Presenter:** DEMCHENKO, Yuri (University of Amsterdam)

**Session Classification:** Demand Of Data Science Skills & Competences