Developing Data Science and Analytics related competences and professional skills:
Practical recommendations to Data Science job seekers, educator and team managers

Introduction and Agenda

Yuri Demchenko, EDISON Project

University of Amsterdam

ICDATA19

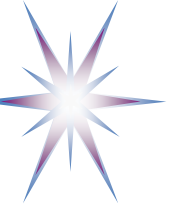30 July 2019, Las Vegas

EDISON
building the data
science profession

FAIRSFAIR
Fostering Fair Data Practices in Europe

mates

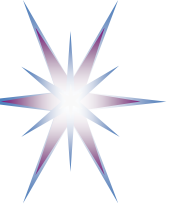# Workshop Agenda

Part 1

1. Introduction and bootstrapping (10 min)
   – EDISON Data Science Framework and EDISON Community Initiative
   – Workshop goals and attendees interests
2. Background information and how the EDSF has been done (15 min)
   – EU and International studies on data related competences and skills
   – EU standards, projects, initiatives, associations
3. EDISON Data Science Framework (EDSF) in details and walk through customised curriculum design (45 min)
   – EDSF components: Competence Framework (CF-DS), Body of Knowledge (DS-BoK), Model Curriculum (MC-DS), Data Science Professional Profiles (DSPP)

Part 2

4. Walk through customised curriculum design: from target professional groups or competences to curriculum suggestions
5. Transversal competences and skills and supporting courses
   – Data Science soft skills: workplace and 21st Century skills
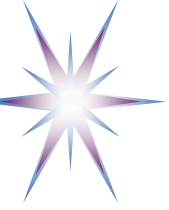6. Closing: questions, discussion

Optional: Data Scientist and Data Steward vacancy and CV analysis and matching from indeed.com – Interactive session and practice (in groups)
• Analysis of Data Science and Data Steward CVs and Job vacancies
• Matching exercise

Data Science Professional Education and Training

Yuri Demchenko, Senior Researcher, Lecturer, UvA

- Graduated and PhD from National Technical University of Ukraine "Kiev Polytechnic Institute"
    - University of Amsterdam – since 2003
- Research areas
    - Big Data Infrastructure and Data Science platforms
    - Intercloud Architecture Framework and cloud powered design, including DevOps and cloud automation platform
    - Cloud security and compliance
- Teaching courses (on campus and online)
    - Big Data Infrastructure and Technologies
    - Data Science Foundations, Professional Issues in Data Science
    - Cloud Computing Engineering, DevOps
    - Security Engineering
- Recent projects
    - EDISON: Building the Data Science Profession for Europe
    - MATES: Digitalisation of the European Blue Economy
    - CYCLONE: Multi-cloud automation platform for cloud based applications
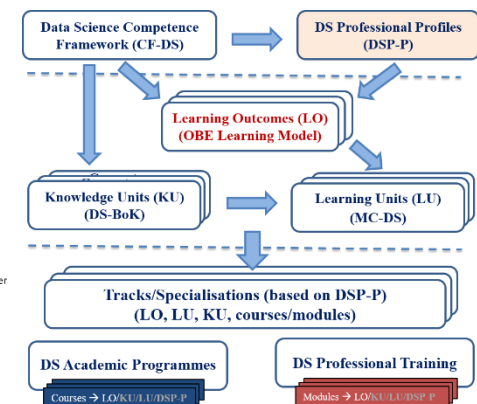   - GEANT4 Research: Cloud aware networking infrastructure provisioning on-demand
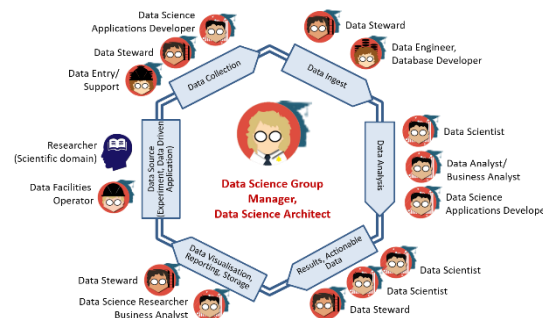
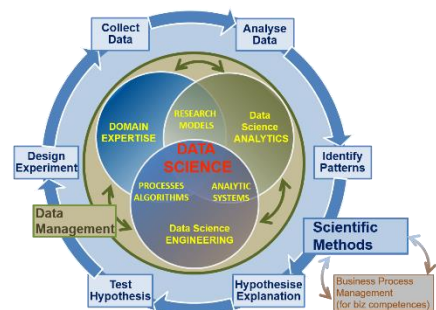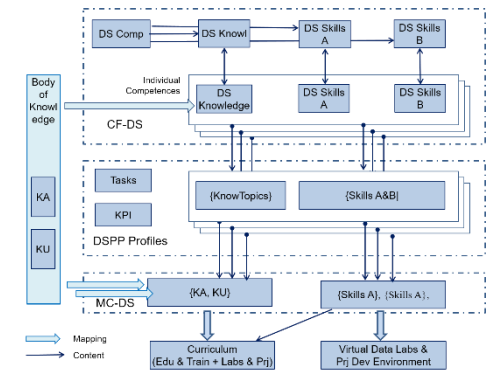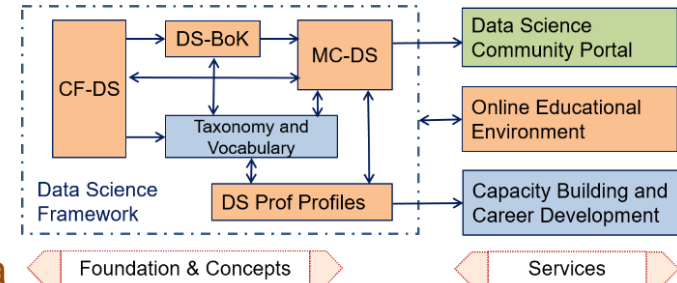Data Science Professional Education and Training
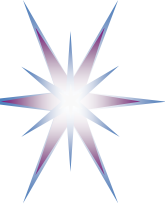
INFRASUPP04 CSA Project EDISON (2015-2017)

- University of Amsterdam (UvA) - NL
- University of Stavanger (UiS) - NO
- University of Southampton (Southampton) - UK
- European Grid Initiative (EGI) - EU
- Engineering (ENG) - IT
- Forschunginstitut for Telekommunikation und Kooperation (FTK) – DE
- InMark - ES

- Expert and professional community support and involvement
  - EDISON Liaison Groups (universities, industry, research)
  - Champion universities
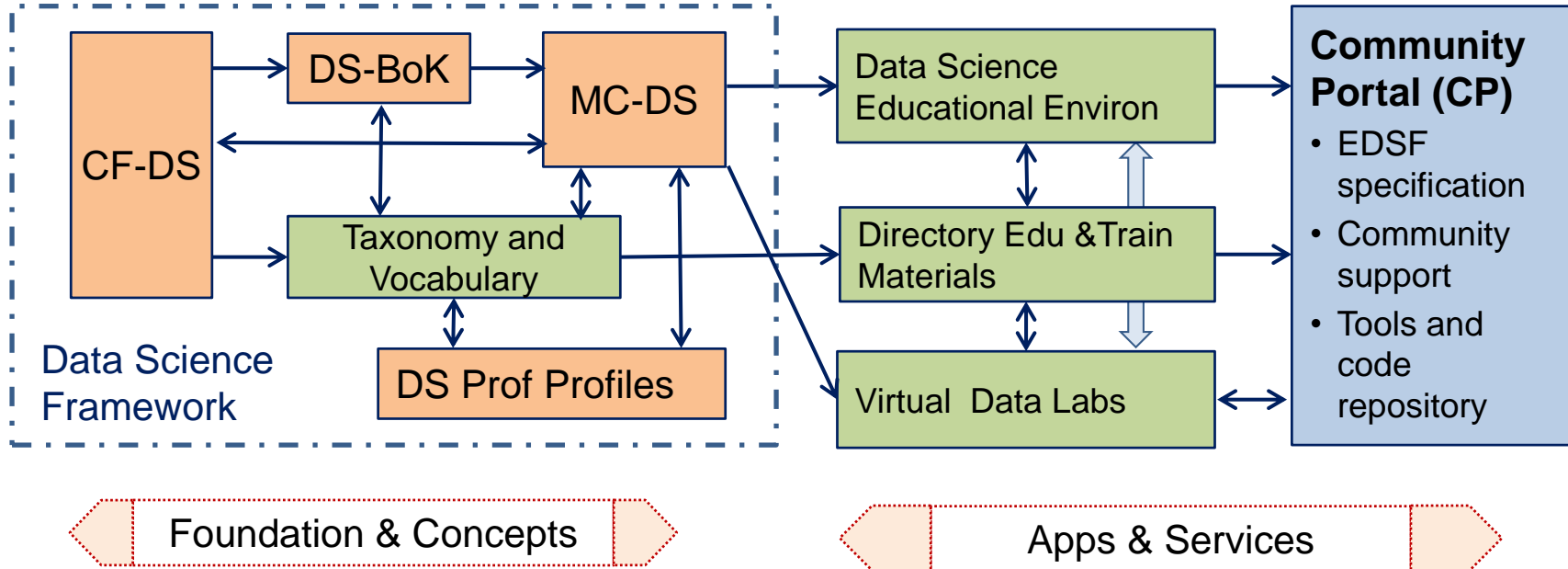 - Research Data Alliance, EUA, CEN, ICTP, BHEF (US), DARE

Data Science Professional Education and Training

EDISON Products for Data Science Skills Management and Curriculum Design

- EDISON Data Science Framework (EDSF)
  - Release 3 components CF-DS, DS-BoK, MC-DS, DSPP
  - Compliant with EU standards on competences and professional occupations e-CFv3.0, ESCO

- Skills development and career management for Core Data Experts and related data handling professions

- Academic programmes and professional training courses (self) assessment and design

- Individual competences benchmarking and Data Science team design

- Cooperation with International professional organisations IEEE, ACM, BHEF, APEC (AP Economic Cooperation)

Data Science Professional Education and Training

EDISON Data Science Framework (EDSF) – Core components and community maintained services
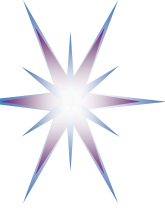

EDISON Framework core components and documents
- CF-DS – Data Science Competence Framework (Part 1)
- DS-BoK – Data Science Body of Knowledge (Part 2)
- MC-DS – Data Science Model Curriculum (Part 3)
- DSPP – Data Science Professional profiles (Part 4)
- Data Science Taxonomies and Scientific Disciplines Classification
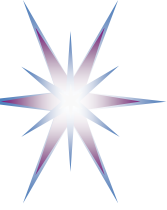
Applications and Services
- Virtual Data Science Labs
- Data Science Educational Environment
- Directory of edu & train resources
- Community Portal – currently github

Data Science Professional Education and Training

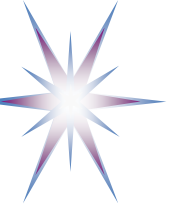# EDSF Recognition, Endorsement and Implementation

- **DARE (Data Analytics Rising Employment)** project by APEC (Asia Pacific Economic Cooperation)
  - DARE project Advisory Council meeting 4-5 May 2017, Singapore

- **PcW and BHEF Report "Investing in America's data science and analytics talent"** April 2017
  - Quotes EDSF and Amsterdam School of Data Science

- **Dutch Ministry of Education recommended EDSF** as a basis for university curricula on Data Science
  - Workshop "Be Prepared for Big Data in the Cloud: Dutch Initiatives for personalized medicine and health research & toward a national action programme for data science training", Amsterdam 28 June 2016
  - Dutch Gov Program on re-skilling IT/data workers for DSA competences

- **European Champion Universities network**
  - 1st Conference (13-14 July, UK)
  - 2nd Conference (14-15 March, Madrid, Spain)
  - 3rd Conference 19-20 June 2017, Warsaw

Data Science Professional Education and Training

Ongoing activities and developments
https://github.com/EDISONcommunity/EDSF
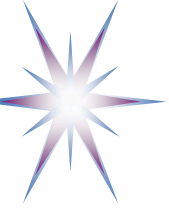
- **EDISON Community Initiative and Call for contribution EDSF Release 4**
  - Call for contribution Part 5 Use ases and Guidelines - Deadline 30 Sept 2019
  - EDSF Design Workshop – 20 November 2019
  - Target publication EDSF Release 4 – December 2019
  - Call for sponsorship

- **Industry digitalisation projects and data literacy skills training**
  - **MATES project funded by EU ERASMUS Programme -** EU Maritime industry digital transformation, data skills development and Data + Ocean literacy
  - Skills for SME: Data Science, IoT, Cybersecurity: Prepare to Data Economy and Industry 4.0

- **DARE project by APEC** (Asia Pacific Economic Cooperation)
  - Continuing cooperation for Asia Pacific region
  - Recommended Data Science and Analytics Competences published August 2017 - https://www.apec.org/Press/Features/2017/0620_DSA
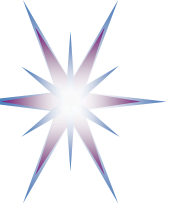
Further development and exploitation

- Development of EDSF based tools
  - Competences benchmarking and Job market monitoring

- Inter-universities initiative "Data Science for UN's Sustainable Development Goals" to focus in-curricula research (projects) on UN priority goals

- Toward EDSF and Data Science profession standardisation
  - ESCO (European Skills, Competences and Occupations) taxonomy
  - CEN TC428 (European std body) – Extending current eCFv3.0 and ICT profiles towards e-CF4 with Data Science related competences
  - Work with the IEEE and ACM curriculum workshop to define Data Science Curriculum and extend current CCS2012 (Classification Computer Science 2012)
  - IBM and OSI initiative to develop Data Science Body of Knowledge

- Data Science Manifesto – Primarily focused on professional and ethical issues in Data Science, new type of professional
  - Professional and ethical issues as a primary focus

Data Science Professional Education and Training

EDISON Initiative Online Presence

- EDSF github project - https://github.com/EDISONcommunity/EDSF
    - Component documents CF-DS, DS-BoK, MC-DS, DSPP

- EDISON Community work area and discussions - https://github.com/EDISONcommunity/EDSF/wiki/EDSFhome
- Mailing list - edison-net@list.uva.nl

- EDISON project website - old domain *edison-project.eu* expired: Legacy information to be moved to http://edison-project.net/

Data Science Professional Education and Training

Links to EDISON Resources

- EDISON Data Science Framework Release 3 (EDSF)
  https://github.com/EDISONcommunity/EDSF

Component EDSF documents

CF-DS – Data Science Competence Framework
https://github.com/EDISONcommunity/EDSF/blob/master/EDISON_CF-DS-release3-v09.pdf
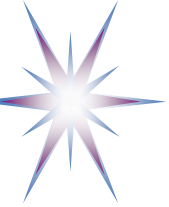
DS-BoK – Data Science Body of Knowledge
https://github.com/EDISONcommunity/EDSF/blob/master/EDISON_DS-BoK-release3-v04.pdf

MC-DS – Data Science Model Curriculum
https://github.com/EDISONcommunity/EDSF/blob/master/EDISON_MC-DS-release3-v04.pdf

DSPP – Data Science Professional profiles
https://github.com/EDISONcommunity/EDSF/blob/master/EDISON_DSPP-release3-v05.pdf

Data Science Professional Education and Training

Other related links

- Amsterdam School of Data Science
  - https://www.schoolofdatascience.amsterdam/
  - https://www.schoolofdatascience.amsterdam/education/

- Research Data Alliance interest Group on Education and Training on Handling of Research Data (IG-ETHRD)
  - https://www.rd-alliance.org/groups/education-and-training-handling-research-data.html