



Data Science among other Data Driven Technology Domains Revisited

Data Science In the Digital Data Driven Economy Ecosystem

- Data Science integrates multiple components from different scientific and technology domains to drive data intensive research and emerging digital technologies development.
- Data Science has emerged as a scientific discipline to become a foundation for academic research and curricula development

Data Science is a complex discipline that uses conceptual and mathematical abstractions and models, statistical methods, together with modern computational tools to obtain knowledge/derive insight from data to uncover correlations and causations in business data and support decision making in scientific research and business activity.
- Data Science uses Scientific/Research Methods for discovery of a value from data: new relations, new factors, knowledge for decision support

Data Scientist definition by the EDISON Data Science Framework (EDSF) – Ref. EDSF, 2016, based on NIST SP 1500

A Data Scientist is a practitioner who has sufficient competences and knowledge in the overlapping regimes of expertise in data analytics skills, domain knowledge, business needs, and programming and systems engineering expertise to manage the end-to-end scientific method process through each stage in the big data lifecycle, till the delivery of an expected scientific and business value to organization or project.

Data Science and Data Handling Professions Family – Compliant with ESCO

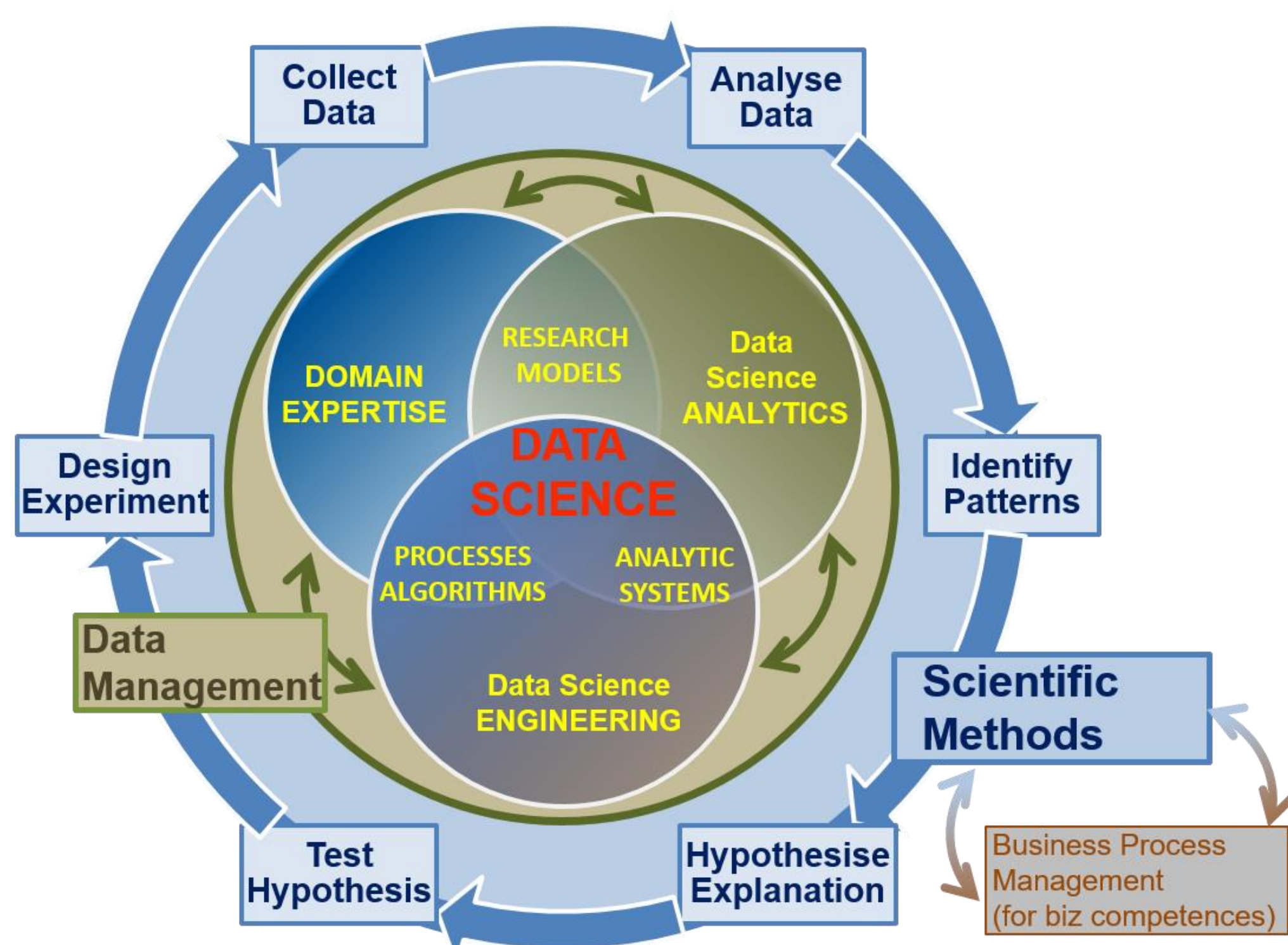
Managers: Chief Data Officer (CDO), Data Science (group/dept) manager, Science infrastructure manager, Research Infrastructure manager

Professionals: Data Scientist, Data Science Researcher, Data Science Architect, Data Science (applications) programmer/engineer, Data Analyst, Business Analyst, etc.

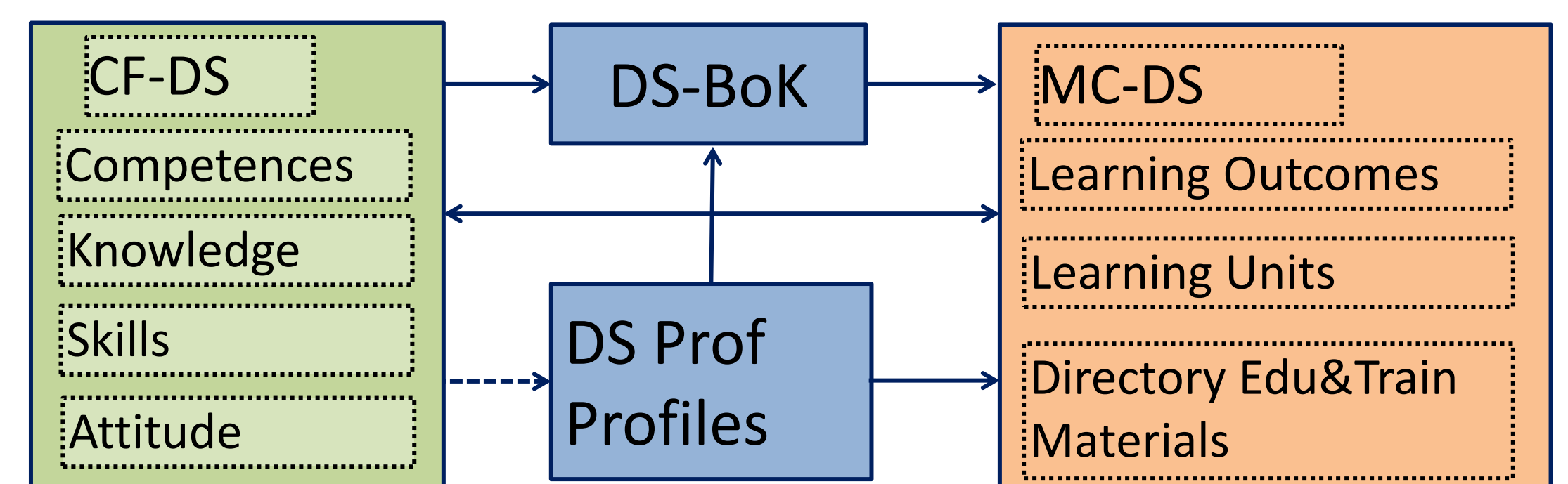
Professional (database): Large scale (cloud) database designers and administrators, scientific database designers and administrators

Professional (data handling/management): Data Stewards, Digital Data Librarian, Data Librarians, Data Archivists

Data Science Competence Domains



EDISON Data Science Framework (EDSF)



Data Science Core Competences and Skills

Core Data Science competences

- Data Science Analytics** (including Statistical Analysis, Machine Learning, Business Analytics)
- Data Science Engineering** (including Software and Applications Engineering, Data Warehousing, Big Data Infrastructure and Tools)
- Domain Knowledge and Expertise** (Subject/Scientific domain related)
- Data Management, Data Governance, Stewardship, Curation, Preservation**
- Research Methods and/vs Business Processes/Operations**

Data Science professional skills: Thinking and acting like Data Scientist – required to successfully develop as a Data Scientist and work in Data Science teams

EDISON Data Science Framework (EDSF) provides conceptual basis for the Data Science profession

- CF-DS – Data Science Competence Framework
- DS-BoK – Data Science Body of Knowledge
- MC-DS – Data Science Model Curriculum
- DSPP - Data Science Professional Profiles
- Data Science Taxonomy and Scientific Disciplines Classification

For more information refer to EDSF Release 3 documents

EDSF github project - <https://github.com/EDISONcommunity/EDSF>

EDISON Community work area and discussions -

<https://github.com/EDISONcommunity/EDSF/wiki/EDSFHome>

Mailing list - edison-net@list.uva.nl

Related publications

The Data Science Framework, A View from the EDISON Project, Editors Juan J. Cuadrado-Gallego, Yuri Demchenko, Springer Nature Switzerland AG 2020, ISBN 978-3-030-51022-0, ISBN 978-3-030-51023-7 (eBook, printed book)

Yuri Demchenko, et al, Data Scientist Professional Revisited: Competences Definition and Assessment, Professional Development and Education Path Design, International Conference on Big Data and Education (ICBDE2021), February 3-5, 2021, London, UK



EDSF is maintained by the EDISON Community Initiative

<https://github.com/EDISONcommunity/EDSF>

Contact: Yuri Demchenko <y.demchenko@uva.nl>

