



Hochschule für  
Wirtschaft und Recht Berlin  
Berlin School of Economics and Law

Conference

# Data Science Framework

## Trust Aspects

05/07/2021, Hochschule für Wirtschaft und Recht Berlin

Berlin, Germany

Prof. Dr. Juan J. Cuadrado-Gallego



Universidad de Alcalá



# Data Science Framework

For a general audience the well-known term is **Big Data** because it can be read or listened to many times each day in all kind of media: Internet, Television, Radio,...

And that general audience, at most know about Big Data what is said about the term in Wikipedia: “Big data is a field that treats ways to analyze, systematically extract information from, or otherwise deal with data sets that are too large or complex to be dealt with by traditional data-processing application software.”

That definition is not wrong, but it is not enough when the approximation to this subject is done in a professional way. The first difference is that the popular term Big Data changes to

## Data Science

Which is the science that underlies the treatment of large data bases.

As we will see in this presentation Data Science is a discipline that **is currently in development and its scope and contents are not closed**, but nowadays it can be defined as:

“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) support decision making in scientific research and business activity”

And, if we must define data science in only one sentence:

**“Science that studies how to obtain knowledge from Data.”**

# Data Science Framework

EDISON Project was the European Union funded Horizon 2020 project with the Grant 675419, which was developed since 2015 until 2017, and its goal was

**to create the foundation for the data science profession for Europe.**

EDISON project undertook multiple initiatives and organized multiple community activities/events and conducted important studies to

**involve data experts and practitioners from academia, research and industry**

**to define the foundation of the new profession of the Data Scientist.**

EDISON project created

**the Data Science Framework (EDSF) repository,**

which is maintained by the EDISON EDSF Community Initiative which involves project partners and **new contributors** (open call for contributors). Updates and revisions are done biannually.

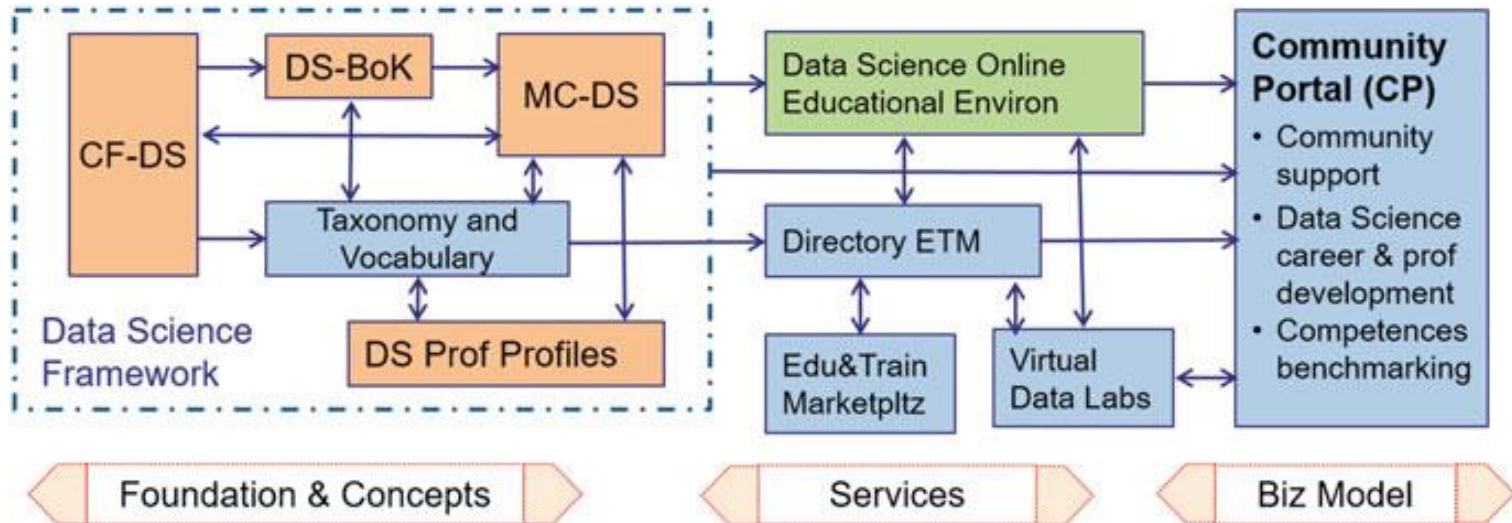
EDISON Project main outcome, which its main results, have been the book:

J.J. Cuadrado-Gallego and Y. Demchenko, Eds., **The Data Science Framework (A View from the EDISON Project)**. Cham, Switzerland: **Springer** International Publishing, 2020. 2020. doi: 10.1007/978-3-030-51023-7. ISBN: 978-3-030-51022-0. pp: XIV, 194. Springer eBook. eBook ISBN: 978-3-030-51023-7

# Data Science Framework

Data Science Framework four corners:

1. Data Science Competence Framework (CF-DS). EDSF Part 1.
  2. Data Science Body of Knowledge (DS-BoK). EDSF Part 2.
  3. Data Science Model Curriculum (MC-DS). EDSF Part 3.
  4. Data Science Professional Profiles and Occupations taxonomy (DSPP). EDSF Part 4.
- Data Science Taxonomy and Scientific Disciplines Classification

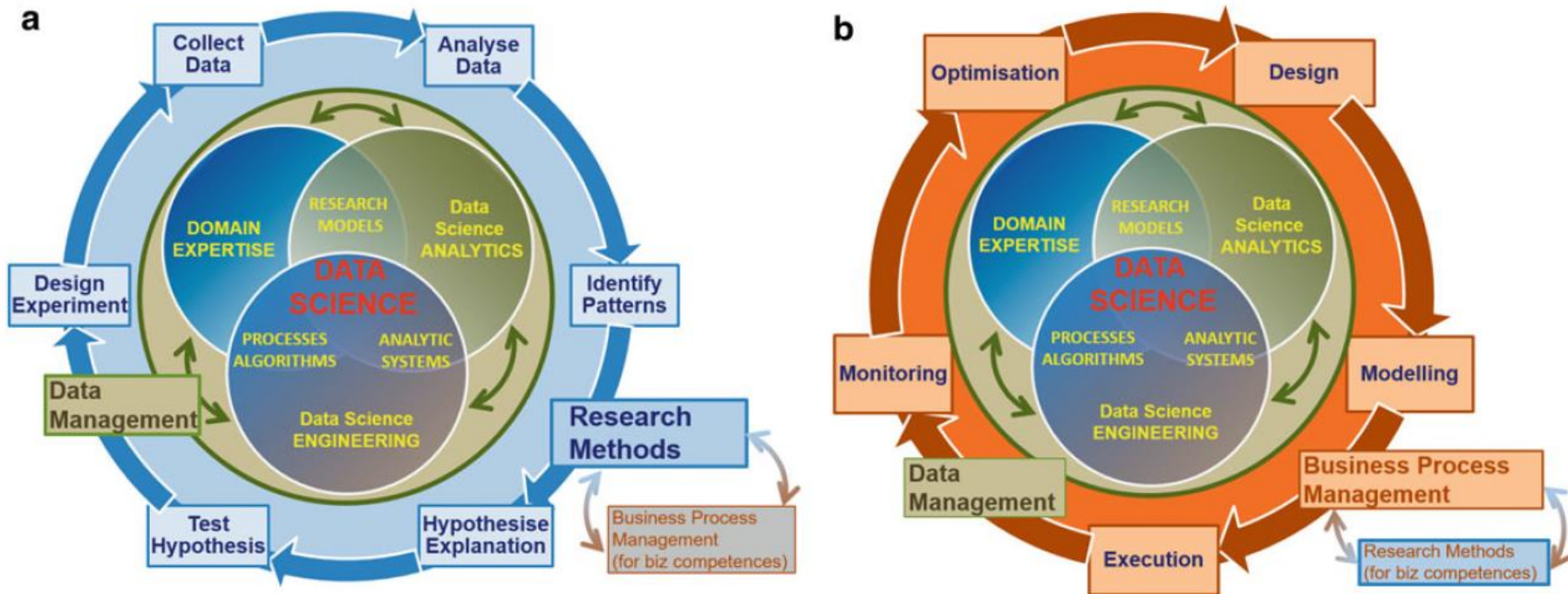


# Data Science Framework

Data Science Competence Framework (CF-DS). EDSF Part 1.

CF-DS adopts a holistic e-CF definition: Competence is a demonstrated ability to apply knowledge, skills and attributes for achieving desirable results in organisational or role context.

Prof. Dr. Juan J. Cuadrado-Gallego



Relations between identified data science competence groups. a. Data science competences for general or research-oriented profiles. b. Data science competences for business-oriented profiles.

# Data Science Framework

Data Science Body of Knowledge (DS-BoK). EDSF Part 2.

## 1. Data Analytics Knowledge Area Group, KAG<sub>1</sub>-DSDA

KA<sub>01.01</sub> (DSDA.01/SMA) Statistical methods for data analysis

KU<sub>1.01.00</sub>. General overview and main concepts in statistical methods for data analysis

KU<sub>1.01.01</sub>. Probability and statistics

... 17 KU.

KA<sub>01.02</sub> (DSDA.02/ML) Machine learning

... 6 KA

## 2. Data Engineering Knowledge Area Group, KAG<sub>2</sub>-DSENG

## 3. Data Management Knowledge Area Group, KAG<sub>3</sub>-DSDM

## 4. Research Methods and Project Management Knowledge Area Group, KAG<sub>4</sub>-DSRMP

## 5. Business Analytics Knowledge Area Group, KAG<sub>5</sub>-DSBA

KA<sub>05.01</sub> (DSBA.01/BAF) Business analytics foundation

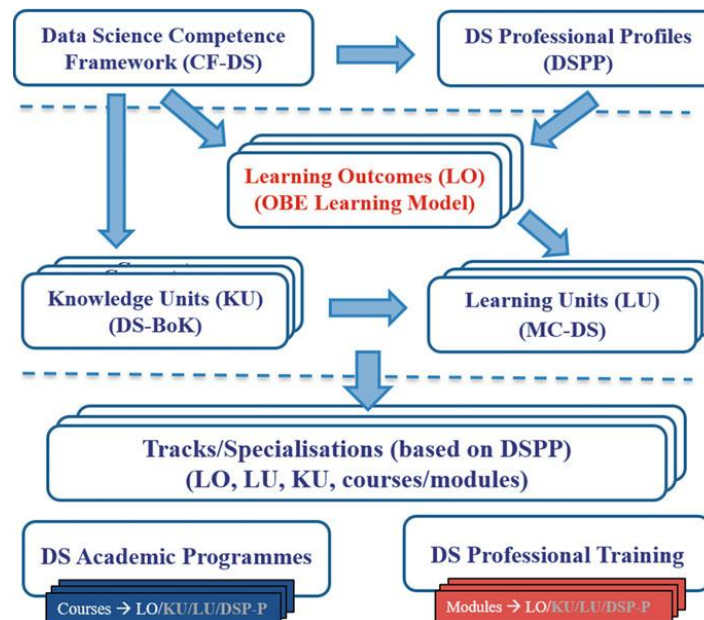
KA<sub>05.02</sub> (DSBA.02/BAEM) Business analytics organisation and enterprise management

# Data Science Framework

Data Science Model Curriculum (MC-DS). EDSF Part 3.

Business Process Management Learning Outcome, LO5-BA: The global learning outcome (LO) of business process management is: Use domain knowledge (scientific or business) to develop relevant data analytics applications and adopt general data science methods to domain specific data types and presentations, data and process models, organizational roles and relations.

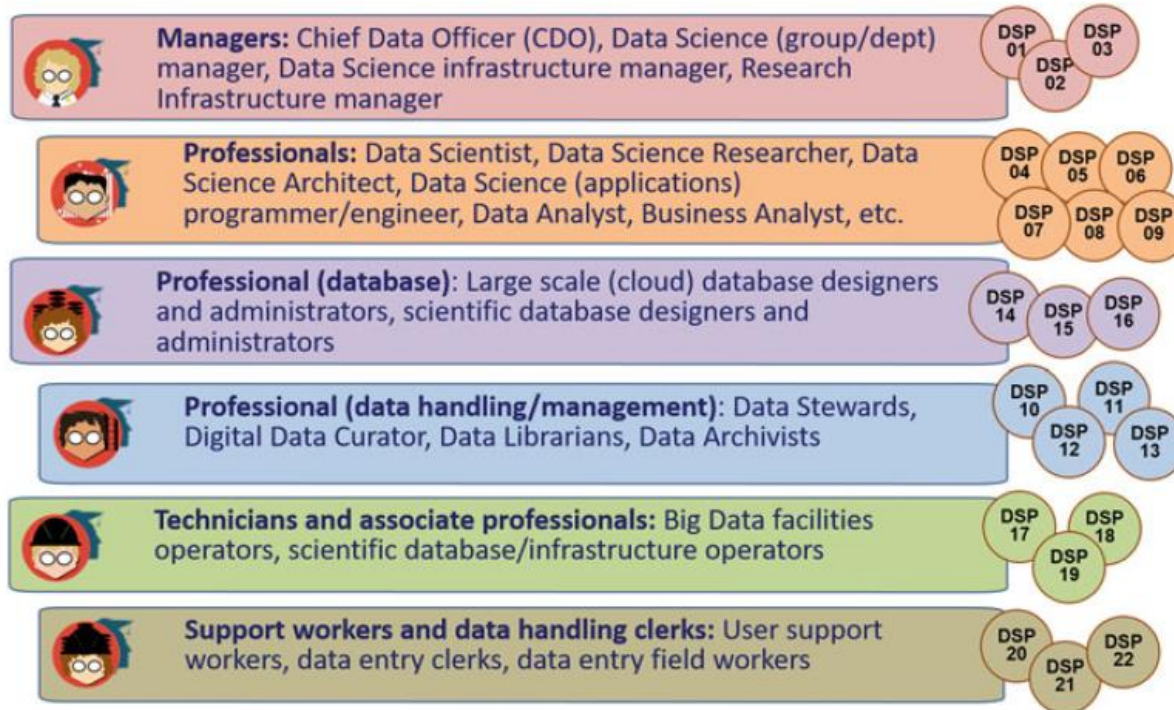
LO for specific DSBA competences are: LO5.01 based on DSBA01. Analyse information needs, assess existing data and suggest/identify new data required for specific business context to achieve organisational goal, including using social network and open data.



Interaction between EDSF components for defining academic or professional curriculum.

# Data Science Framework

Data Science Professional Profiles and Occupations taxonomy (DSPP). EDSF Part 4.



Data Science 22 Professional Profiles and their grouping by the proposed new professional groups.



# Data Science Framework

To know more about the Data Science Framework you can go to:

<https://www.springer.com/gp/book/9783030510220#>

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

And if you want to contact me:

Prof. Dr. Juan J. Cuadrado-Gallego

[jjcg@uah.es](mailto:jjcg@uah.es)



Hochschule für  
Wirtschaft und Recht Berlin  
Berlin School of Economics and Law

# Thank you!

Conference

## Data Science Framework Trust Aspects

05/07/2021, Hochschule für Wirtschaft und Recht Berlin

Berlin, Germany

Prof. Dr. Juan J. Cuadrado-Gallego



Universidad de Alcalá

