
Data Science

Fields of education: Engineering and information technology

1. Professional qualification

Professional career outline

Professional data scientists are contributing to the design, development and deployment of so-called data products or data pipelines in enterprises and public institutions. The analysis of current occupational profiles in Switzerland¹ is showing that Data Science involves three pillars of activities. The first one can be summarised by the term data engineering that involves the collection, preparation, storage and processing of enterprise data using classical database tools and more and more often scalable big data infra-structure. The second area of activities is data analytics that englobes the use of statistical methods and machine learning including deep learning and natural language processing to build models to describe and learn from data or to support business decisions. The third area can be called data services and covers the organisation, the deployment and management of data pipelines involved in data products and services to solve real-life business problems while meeting requirements in terms of service level, data security, ethical and legal aspects. Figure 1 below illustrates the positioning of the Data Science profile at the inter-section of these three pillars.

Professional skills

Students graduating as "MSE Data Scientists" will have developed interdisciplinary professional skills that can be described according to the three areas of activities of Figure 1:

Data Analytics

- Applying statistical methods to describe and explore data as well as to draw conclusions from data subject to random variations.
- Using data mining, machine learning and deep learning to build data-driven models.
- Planning and conceptualizing Design of Experiments that follow state-of-the-art procedures for building and evaluating usable models and interpretable results.

Data Engineering

- Organising the collection and sourcing of application-related data from heterogeneous origins such as Industrial IoT sensors, structured relational databases, unstructured distributed data, big data stacks and online streaming data feeds.
- Planning and organising the storage of such data using the adequate tools and resources that meet application constraints such as bandwidth, safety and scalability.
- Developing applications able to process data streams, extract features and apply models on it.

Data Services

- Building data products: understanding business needs in terms of data pipelines able to turn analysis results into business insights and actions.
- Organising the deployment, the maintenance and the evolution of data services respecting service level agreements.
- Understanding non-technical constraints related for example to ethics, privacy, security and safety while processing data in enterprises and public institutions, as well as organizing data services to meet such constraints.

¹ An analysis of job profile descriptions on sites such as www.jobs.ch was performed, as well as interviews of data scientist currently active in the economic sector in Switzerland.

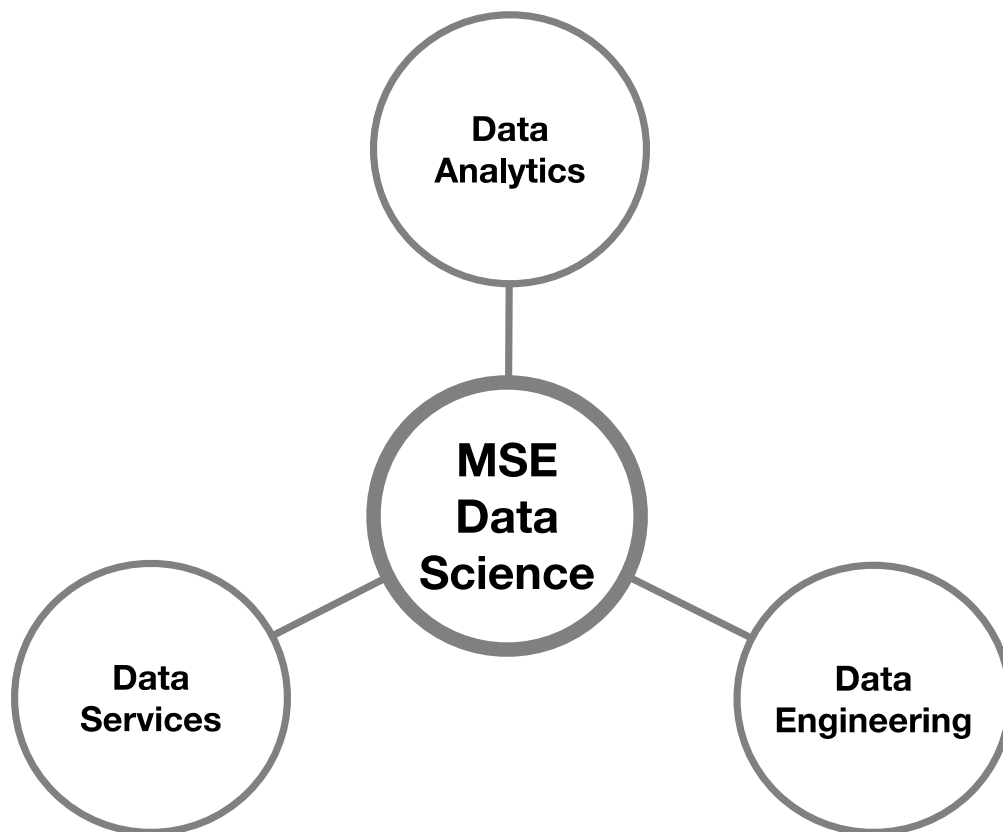


Figure 1. The three pillars of the MSE Data Science Profile

Entry skills

Specific skills are required to enrol in this profile. Interested students will be individually assessed for their suitability by the respective University of Applied Sciences. The assessment of the entry skills is part of the enrolment process of the respective school.

Differentiation to bachelor level

The closest bachelor engineering degree related to Data Science are:

- Degrees oriented towards ICT / Computer Science. Such bachelor degrees are covering topics more directed towards data engineering and algorithmic foundations. The MSE Data Science profile will complete the education towards topics related to Data Analytics and Data Services, including advanced competencies in statistics, data mining, machine learning, deep learning, and towards the design, the implementation and the deployment of data services for enterprises.
- Degrees oriented towards Engineering and Management. Such bachelor degrees are covering topics more directed towards Data Services. Here, the MSE Data Science profile will complete the Data Engineering and Data Analytics parts and their related targeted competencies.

For all the other bachelor-level engineering degrees, the Data Science profile will provide a unique opportunity for students to develop their knowledge in applied Data Science, providing data-oriented competencies on top of their specific engineering knowledge acquired in the bachelor. For example, an engineer with a bachelor degree in chemistry who would be continuing with a Data Science profile would qualify for planning, designing, building and maintaining data pipelines for chemical industries.

2. Profile contents

The profile covers the following content:

Figure 2 below provides a summary of the three areas of education on which the Data Science profile will be articulated. It is expected that these areas are covered by the different forms of modules (CM, FTP, TSM, EVA, deepening modules etc.) but also by projects (deepening and thesis).

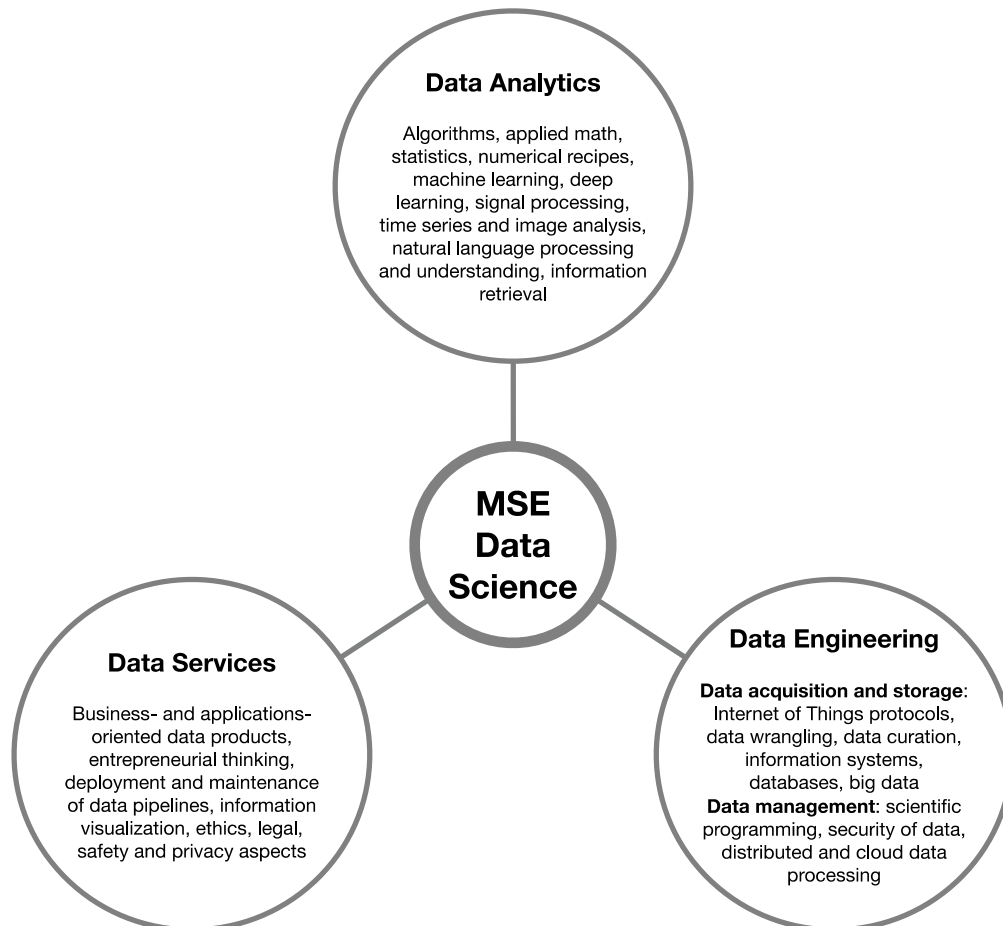


Figure 2 - Summary of covered topics in the DS profile

«Data analytics» is the process of converting raw data into information, models and actionable knowledge. In this process, models and algorithms are applied to data. They are based on applied mathematics, algorithmic, analysis of data such as time series and images, information retrieval, machine and deep learning, natural language processing and understanding, signal processing, statistics, visual analytics, etc.

«Data engineering» deals with data acquisition, storage and processing from a computer science perspective. In the context of data acquisition and storage, data engineering includes data sensing protocols such as the ones used for the Internet of Things, the communication with databases and data streams, as well as the use of scalable data storage systems for big data. Data processing involves the scientific programming for data wrangling, data security and distributed processing.

«Data services» is about making data actionable from a business and service perspective. This requires a serious understanding of the underlying processes and tasks, how such data-driven products and pipelines are designed and operated, how information can be visualised and communicated, how algorithms and models are maintained during their life-cycle, etc. Finally and importantly, ethical, legal, safety and privacy issues need to be considered in the deployment and maintenance of such data-driven services.