



MONTPELLIER



MAJOR

Data Engineering

#RESPONSIBLE DATA SCIENCE

#ARTIFICIAL INTELLIGENCE

#BIG DATA

#COLLABORATIVE DEVELOPMENT



ENGINEERING SCHOOL
Creating the future together

PROGRAM AIMS

The aim of the **Data Engineering** major is to train flexible and adaptable engineers both strong in Mathematics & Computer Sciences. They will be able to **help companies and laboratories to structure their data** and to **produce new insights** with Data analysis and Machine Learning approaches. Emphasis is placed on a systemic approach (cost/benefit) including **legal, human, economic and environmental aspects**.

Combining their **generalist education** with **an expertise on all the Data Journey**, the graduates will be able to address all the missions of the **Data Analyst**: data recovery, structuration, analysis & reporting, in direct interaction with the end-user but will also have a strong suit in Data Science allowing them to create models and advanced tools (predicter, classifier,...) and sufficient knowledge on the **Data Engineering** side to collaborate on the actual transformation of those models into services.

With a pedagogical approach **based on skills, learning-by-doing and life-long learning** the graduates should be able to integrate small and big structures and to adapt to many different business..



PROGRAM STRUCTURE

The major extends over two academic years and is organised around two inclass semesters, framed by two internship semesters. (Note : For the international students, the first internship is replaced by an International Project semester which includes mechanics, energy, computer science and French.)

All the CUs are offered **in English**. They are designed as independent credits so as to admit students from other programs or students attending vocational training.

In order to be as close as possible to employment conditions, the Major's project CU use a project approach, thus confronting students to a real client specifications, teamwork and autonomy.

YEAR 4 . COMPULSORY CUS

Computer Science 101 	69 hours 5 ECTS
<ul style="list-style-type: none">· Infrastructure (OS & Virtualization)· Development (Web programming & Collaborative development)	Hardware & software basics for future computer science's engineers.
Data architecture 	69 hours 5 ECTS
<ul style="list-style-type: none">· Data Models (Relational & Non-relational)· Data Processing Development (Parallelization, Distributed computing)	Data architecture from client's need to physical implementation and considerations on how to process large quantity of data.
Exploratory Data Analysis 	75 hours 5 ECTS
<ul style="list-style-type: none">· Basics of Probability & Statistics· Data Manipulation & Visualisation (Panda & co)	The mathematical and software tools to explore the data and produce insight for the clients.
Basics of Machine Learning 	69 hours 5 ECTS
<ul style="list-style-type: none">· Maths for Machine Learning (Analysis, Algebra, ...)· Data Cleaning· Introduction to Machine Learning	The maths behind the magic. What are the types of problem that can be solved by Machine Learning. Starting to teach the computer how to play with data.
Support Digital Transformation 	64 hours 5 ECTS
<ul style="list-style-type: none">· Digital Transformation· Change management· Professional english	See the big picture of Digital Transformation and how to accompany it.
Major's project 	150 heures 5 ECTS
<ul style="list-style-type: none">· Project management· Learning-by-doing semester's project	Theory and practical application of AGILE methodology.



YEAR 5 . COMPULSORY CUS

Data diversity 	63 hours 5 ECTS
<ul style="list-style-type: none">· Data Sources (IoT, Surveys, Digital traces & API)· Geographical data (GIS & Remote sensing)	An exhaustive panorama of the sources of Data and a focus on less common ones : maps & satellite imagery.
Machine Learning : Theory & Practice 	72 hours 5 ECTS
<ul style="list-style-type: none">· Predictive modeling & clustering· Natural Language Processing· From PoC to production	Everything you always wanted to know about ML and its implementation in the real word. Focus on the specificity of Natural Language Processing.
Responsible Data Science 	66 hours 5 ECTS
<ul style="list-style-type: none">· Data Ethics· Data Law· Ecological impact of IT	Real world issues : take a step back from technology.
Data Strategy 	69 hours 5 ECTS
<ul style="list-style-type: none">· Data for Business Intelligence· Data Governance· Data Protection (Cybersecurity, Block-chain)· Data Science Challenge	How the company deals with the new data-centric paradigm and how the future Data engineer can contribute to it.
Major's project 	150 hours 5 ECTS
<ul style="list-style-type: none">· Project Management· Professionalization· Project monitoring & final presentation	Team project on behalf of a client.

YEAR 5 . ELECTIVE CUS*

Deeper into Data Science 	66 hours 5 ECTS
<ul style="list-style-type: none">· Foundation of Machine Learning· Dimensionnality Reduction· Times Series· Computer Vision with ML	Mathematical explanation on why ML works. How to cope with large variety of data. Specific math models for time series. Image classification : for fun & profit !
Deeper into Data Engineering 	66 hours 5 ECTS
<ul style="list-style-type: none">· Pipelines & Data flow· Big Data· Cloud-based production	Engineering of the Data Journey. How to cope with very large database. How to setup a cloud-based data pipeline.

* One of the two ELECTIVE CUS must be chosen. Opening of the elective is subject to a minimal number of student.

YOUR FUTURE AFTER THE DATA ENGINEERING MAJOR :

The array of business sectors of interest is extremely wide and concerns large industrial groups as much as SME/SMIs, research centres or start-ups.

TARGETED POSITIONS

- Data Analyst
- Data Scientist
- ML Data Engineer

And many other position in the IT and Data sector.

FIELDS

- Companies in the digital industry
- Insurance and health companies
- Banks/ Financial industry
- Sales, distribution/ Marketing
- Medical/ pharmaceutical industry
- Energy
- Communal services
- Industry
- Transport industry
- Life sciences
- Natural Sciences
- Engineering
- Journalism

PROJECTS

A project is carried out on both academic semester in collaboration with a company. **It is used as a guideline for the whole semester and serves as support to the many lectures.**

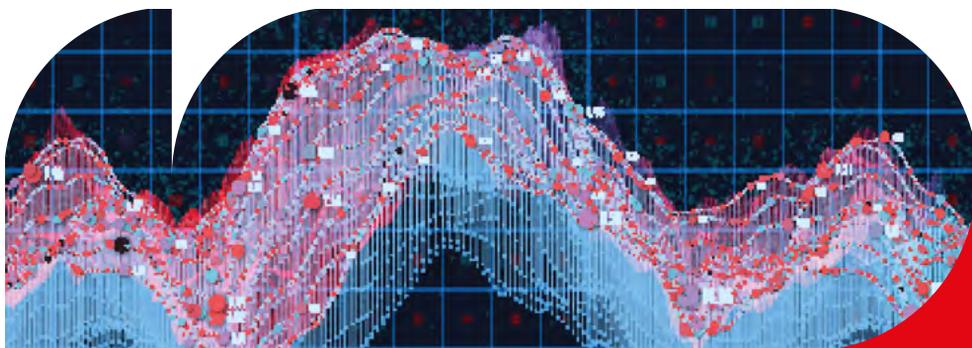
Some examples of projects conducted in the major:

► 4th year:

- Development of a **web tool** allowing the co-design between volunteers and researchers of a common **OLAP-cube representation** of a **biodiversity** database.
- **Data structuration** (model & tools) of a **financial** analysis process in renewable energy sector.

► 5th year:

- Use of **Natural Language Processing** techniques for the **analysis and clustering** of a large **medical's** survey database.
- Development of **Deep Learning** tools for **Object detection** in aerial imagery applied on **ecological study** of cetaceans.



Jade KRESIBERGER

2021 School Year

“ The generalist course of the EPF allowed me to take the time to discover the different fields of engineering. The choice of our studies can be complex. I have always admired people who develop computer tools, but despite my curiosity, I did not know if this area would suit me.

During my first 3 years at EPF in Montpellier, I took my first steps in programming, development or even data analysis. We had strict plans and guidelines to follow while retaining a certain freedom that also allowed us to express our personality. From scratch, we managed to create games or tools for data management, which gave me a lot more and comfort my choices to do the Major in Data Engineering.

We studied a lot of different subjects : data analysis, Machine Learning, Big Data or even Ecodesign. All the things we could do were amazing! From a simple survey, it was possible to extract conclusive information. For example, my 4th year project consisted of carrying out a study on companies that use connected objects for their professional lives. We were able to estimate the sectors that were the fondest about them and why companies were not able to use them (budget...).

The Data Engineering major and the semester project allowed me to discover IT project management, which combines team management and work progress. I had already appreciated it during my internship at Bollor2 Logistics in Singapore. I had the chance to develop a tool for estimating the cost and CO₂ emissions of air and maritime transport for customers in order to reduce the number of company e-mail exchanges. **Yes! IT can be useful for our planet!**

Currently in 5th year I have started a professionalization contract at Accenture, within the Artificial Intelligence / Data department where I will assist a Data Architect. I no longer have any doubts that I belong in this exciting field because I participate in projects which make sense! ”

ANY QUESTIONS?

Antoine GADEMER
Robert RAPADAMNABA
majeure-data@epf.fr

OPEN HOUSE 2021 | 2022, dates, timetables
and registration on www.epf.fr



PARIS-CACHAN CAMPUS . 55 av. du Président Wilson . 94230 Cachan . +33 (0)1 41 13 01 51
TROYES CAMPUS . 2 rue Fernand Sastre . 10430 Rosières-près-Troyes . +33 (0)3 25 70 77 19
MONTPELLIER CAMPUS . 21 bd. Berthelot . 34000 Montpellier . +33 (0)4 99 65 41 81
SAINT-NAZAIRE CAMPUS . 107 av. de la République . 44600 Saint-Nazaire

epf.fr