

Master Data AI: Welcome and Introduction

Goran Frehse & Louis Jachiet

03/09/2021

Overview

DataAI Team

Choosing and Validating Courses

Getting an M2 in Data AI

Getting an M1 in Data AI

DataAI Basics Week

DataAI M1 Example Schedule

DataAI Optional Courses

Table of Contents

DataAI Team

Choosing and Validating Courses

Getting an M2 in Data AI

Getting an M1 in Data AI

DataAI Basics Week

DataAI M1 Example Schedule

DataAI Optional Courses



Louis Jachiet
Coordinator



Goran Frehse
Coordinator



Danielle Deloy
Administrator

Table of Contents

DataAI Team

Choosing and Validating Courses

Getting an M2 in Data AI

Getting an M1 in Data AI

DataAI Basics Week

DataAI M1 Example Schedule

DataAI Optional Courses

Choosing Your Courses

Try to be coherent: For M1, start with basic courses.

- Fill in the web form; saving it checks for alerts.
- Submit schedule by Tuesday, September 7 evening.

Choosing Your Courses

Try to be coherent: For M1, start with basic courses.

- Fill in the web form; saving it checks for alerts.
- Submit schedule by Tuesday, September 7 evening.
- **DEMO**

Validating a Course

- A 24h course = 2.5 ECTS (most Data AI courses).
- A X course = 5 ECTS (more ECTS but also more work).
- A course is validated when grade ≥ 10 out of 20.
- Exams can be hard.
- Failed exam can **SOMETIMES** be retaken in 2nd session.
- **DON'T CHEAT**
 - sanctions are severe,
 - plagiarizing is cheating,
 - if you're desperate, come and see us rather than cheat.

Table of Contents

DataAI Team

Choosing and Validating Courses

Getting an M2 in Data AI

Getting an M1 in Data AI

DataAI Basics Week

DataAI M1 Example Schedule

DataAI Optional Courses

Getting an M2 in Data AI

- 30 ECTS in courses (12 courses), with 25 ECTS from DataAI and 5 ECTS chosen freely
- Validate the course requirements
 - at least one course from each mandatory groups (except those already taken in M1)
- Master's internship (Master's thesis) = 30 ECTS, between Feb. and Sept.

Mandatory Course Groups

One from each group, no overlapping topics (see course descriptions).

- **Group Machine Learning**

- Machine Learning: Shallow & Deep Learning
- Machine & Deep Learning Introduction
- Machine Learning

- **Group Logics**

- Logics and Symbolic AI
- Logic & Knowledge representation

- **Group Databases**

- Database management systems
- Databases

- **Group Big Data Systems**

- Systems for Big Data
- Architectures for Big Data
- Big data infrastructures
- Big Data Processing

- **Group Softskills**

- Softskills seminar

- **Group Ethics**

- AI Ethics

M1/M2 Internships

- Internship in a research lab (very much encouraged) or in a company (discouraged)
- DataAI will post internship on the website. Télécom Paris organizes internship forums.
- Internship advisor supervises the student at host institution
- Scientific advisor (lecturer of the Data AI) mediates in case of disagreement between the student and the internship advisor
- Written report, oral defense, both in English
- Validated if the grade is at least 10/20.

M1/M2 Internships – What To Do Beforehand

1. Fill out the form in the Synapses system (fill in all fields)
2. Send the internship offer by email to the DataAI coordinators for approval
3. Signal scientific advisor to DataAI administrator
4. Fill out the internship agreement (“CONVENTION DE STAGE”)
5. Print 6 (!) copies of the internship agreement and give them to DataAI administrator for signature

M1/M2 Internships – What To Do At the End

1. Send report to scientific advisor 1 week before the defense
2. Defense: 20 minutes presentation, then questions by the advisors
 - no more than 3 weeks before and 1 month after the end of the internship; during the current year of study
 - takes place at the reference institute (Télécom Paris) or at the institute of the scientific advisor
 - attended by scientific advisor and internship advisor if possible
3. Internship is graded by the scientific advisor in coordination with the internship advisor: quality of the work, the report, and the defense

M1 Internships – Specifics

- 2–3 months
- to be completed by the 15th of August
- Report 20 pages, oral defense, both in English

M2 Internships – Specifics

- 5–6 months, starting between Feb. and May.
- to be completed by the 1st of October.
- Report 30-60 pages, oral defense, both in English

Table of Contents

DataAI Team

Choosing and Validating Courses

Getting an M2 in Data AI

Getting an M1 in Data AI

DataAI Basics Week

DataAI M1 Example Schedule

DataAI Optional Courses

Getting an M1 in Data AI

During the M1 year, a student **can** validate:

- up to two research projects awarded 5 ECTS each
- a 2-3 months internship
- 10 ECTS outside of DATA AI courses with at most 5 ECTS in non-math/CS topics
- a certain number of DATA AI courses.

To validate the M1 year, a student **must** validate

- two research projects OR one internship and at least one research project.
- a total of 60 ECTS.

Remember to take some courses from mandatory course groups (required for M2)

M1 Research Projects

- 2 Research projects, each \sim 10 days work scattered throughout a semester
- Requires finding an advisor & approval by Data AI team.
 - both projects can have the same adviser
 - second project approved only after first one is finished
- We will provide a list of research projects topics, but you can also propose your own.
 - re-implementing well known techniques, benchmarking different approaches, reading articles, writing code or ideas, making experiments, etc.
- Report: write a short document summarizing your contribution.
- Poster session at the end of the year
 - The poster can represent one or both projects.

Table of Contents

DataAI Team

Choosing and Validating Courses

Getting an M2 in Data AI

Getting an M1 in Data AI

DataAI Basics Week

DataAI M1 Example Schedule

DataAI Optional Courses

DataAI Basics Week, 6-10 Sept.

| | Monday | Tuesday | Wednesday | Thursday | Friday |
|---------------------|--|---|---|----------|------------------------------------|
| 8:30 – 11:45 | Algorithms and data- structures Angelos Anadiotis | How to use a computer like a computer scientist Louis Jachiet | | | lab exercises |
| 13:30 – 16:45 | Statistics and probability Tiphaine Viard | Formal languages Fabian Suchanek | Logic Jean-Louis Dessalles | | lab exercises OR SD202/IA301 |

Table of Contents

DataAI Team

Choosing and Validating Courses

Getting an M2 in Data AI

Getting an M1 in Data AI

DataAI Basics Week

DataAI M1 Example Schedule

DataAI Optional Courses

Example M1 Schedule, Period 1

| | Monday | Tuesday | Wednesday | Thursday | Friday |
|----|---|---------|--|----------|------------------------|
| AM | DataAI901 Machine Learning: Shallow & Deep Learning | | TPT-SD201 Mining of Large Datasets | | |
| PM | DataAI964 Knowledge Base Construction | | DATAAI922 Big Data Processing | | TPT-SD202 Databases |

Example M1 Schedule, Period 2

| | Monday | Tuesday | Wednesday | Thursday | Friday |
|----|--|------------------------|--|--|---|
| AM | TPT- DATAAI921 Architectures for Big Data | | TPT- DATAAI963 Navigation for autonomous systems | | |
| PM | | DataAI951 AI Ethics | | DATAAI902 Recents Trends in Deep Learning | TSP-IA304 Probabilistic Models and Machine Learning |

Example M1 Schedule, Period 3

| | Monday | Tuesday | Wednesday | Thursday | Friday |
|----|---|--------------------------------------|---|--|--------|
| AM | TPT-IA312 Machine Learning for Text Mining | ENSTA-IA305 Learning for robotics | TPT-SD206 Logic & Knowledge representation | DataAI965 Basics of image processing and analysis | |
| PM | TPT-IA312 Machine Learning for Text Mining | | | DataAI962 Data Stream Mining | |

Example M1 Schedule, Period 4

| | Monday | Tuesday | Wednesday | Thursday | Friday |
|----|--------|---------|-----------|----------|---------------------------|
| AM | | | | | |
| PM | | | | | TPT-SD212 Graph Mining |

You can fill your free-CS ECTS from CS courses at IP Paris. Some examples that you can find in Synapses of **Télécom** (confirm with teacher whether they are in English):

- 2A Track "Traitement du signal pour l'intelligence" (TSIA)
 - SD-TSIA204 : Statistics: linear models
 - SD-TSIA205 : Advanced Statistics
 - SD-TSIA210 : Machine Learning
 - SD-TSIA211 : Optimization for Machine Learning
 - SD-TSIA214 : Machine Learning for Text Mining

You can fill your free-CS ECTS from CS courses at IP Paris. Some examples that you can find in Synapses of **Télécom** (confirm with teacher whether they are in English):

- INF280 : Programming contest training
- DS-telecom-14 - Introduction to Deep Learning with Python part 1
- DS-telecom-21 - Introduction to Deep Learning with Python part 2

You can fill your free-CS ECTS from CS courses at IP Paris. Some examples that you can find in Moodle of **Polytechnique** (confirm with teacher whether they are in English):

- INF634 - Computer Vision
- X-INF643 : Soft Robots : Stimulation, Fabrication and Control
- X-MAP/INF641 : Reinforcement Learning (math-heavy)

Free-non-CS ECTS

You can fill your free-non-CS ECTS from courses at IP Paris that are not related to CS. Some examples that you can find in Synapses of **Télécom** (confirm with teacher whether they are in English):

- MODS214 : Economics
- SES204 : Internet and Society
- for more, go to "Catalogue 2021-2022", then in the menu on the left select "Anglais" under "Langue d'enseignement" (93 courses)
- language courses (FEL)

Table of Contents

DataAI Team

Choosing and Validating Courses

Getting an M2 in Data AI

Getting an M1 in Data AI

DataAI Basics Week

DataAI M1 Example Schedule

DataAI Optional Courses

Probabilistic Models and Machine Learning (TSP-IA304)

Machine Learning in High Dimension (TPT-IA317)

Kernel Machines (TPT-IA326)

Advanced Machine Learning and Autonomous Agents (X-INF581-1)

Recent Trends in Deep Learning (TPT-DATAAI902)

Programming with GPU for Deep Learning (TPT-IA307)

Graph Mining (TPT-SD212)

Graph mining and Clustering (TPT-MITRO209)

Data Stream Mining (TPT-DATAAI962)

Mining of Large Datasets (TPT-SD201)

Image mining and content-based retrieval (TPT-AIC-DK921)

Topological Data Analysis (X-INF556)

Knowledge Base Construction (TPT-DATAA1964)

Semantic Networks (CSC5003-2)

Efficient resolution of logical models (ENSTA-IA303)

Constraint programming (ENSTA-IA302)

Machine Learning for Text Mining (TPT-IA312)

Text Mining and NLP (X-INF582)

Cognitive approach to NLP (TPT-SD-213)

Multimodal Dialogue (TPT-IA315)

Self-Organising Multi-Agent Systems (TPT-DATAAI961)

Navigation for autonomous systems (TPT-DATAAI963)

Learning for robotics (ENSTA - IA305)

Reinforcement Learning (TPT-IA318)

Emergence in Complex Systems (TPT-AthensTPT-09)

Basics of image processing and analysis (TPT-DATAAI965)

Algorithmic information and artificial intelligence (TPT-IA703)

Data Visualization (X-INF552)

Questions?