

Share



Email destinataire :

Vous devez renseigner un email valide

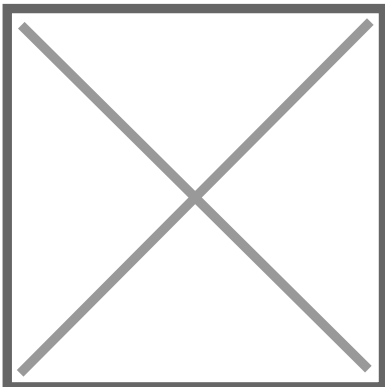
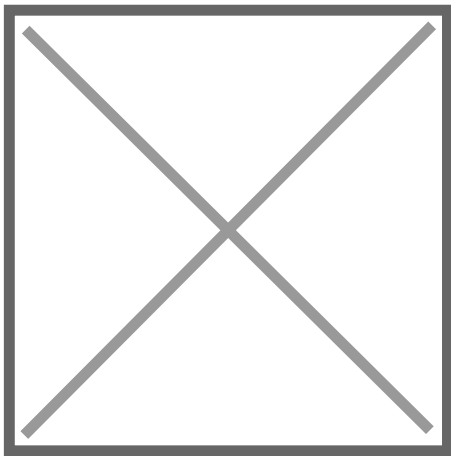
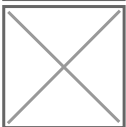
Envoyer

Une erreur est survenue. Merci de réessayer ultérieurement

Le mail de partage a bien été envoyé.

Add to Compare

M2 Data management in large-scale distributed systems



Master's degree

Specialisation Computer Science

Enrolment type

Full-time academic programmes

Language(s) of instruction

French

The **DataScale** program trains experts capable of mastering the entire data value chain: data engineering, architectures and advanced AI analytics. It leads to diverse careers (data scientist, data engineer, AI analyst, etc.) across key sectors, and also provides access to doctoral studies in public or private laboratories.

The program is offered as an initial training course. Classes take place from mid-September to the end of March, followed by a 5- to 6-month internship starting in April.

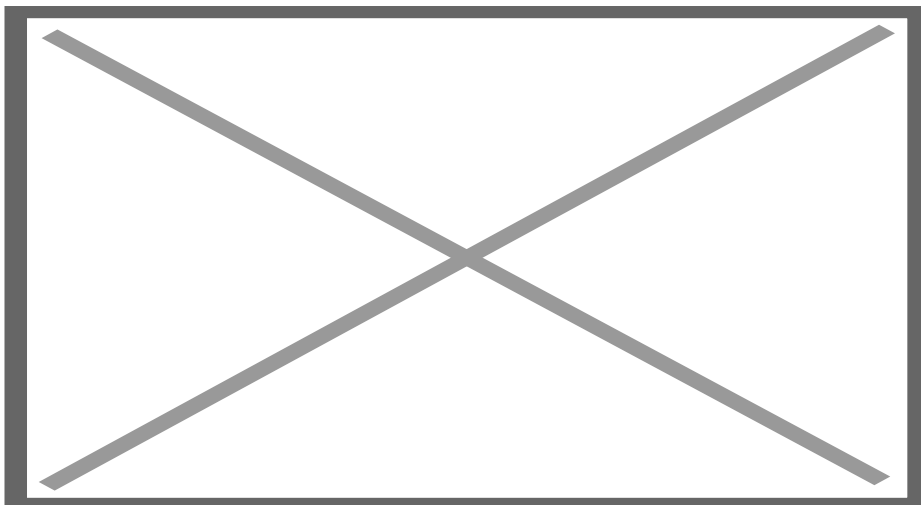
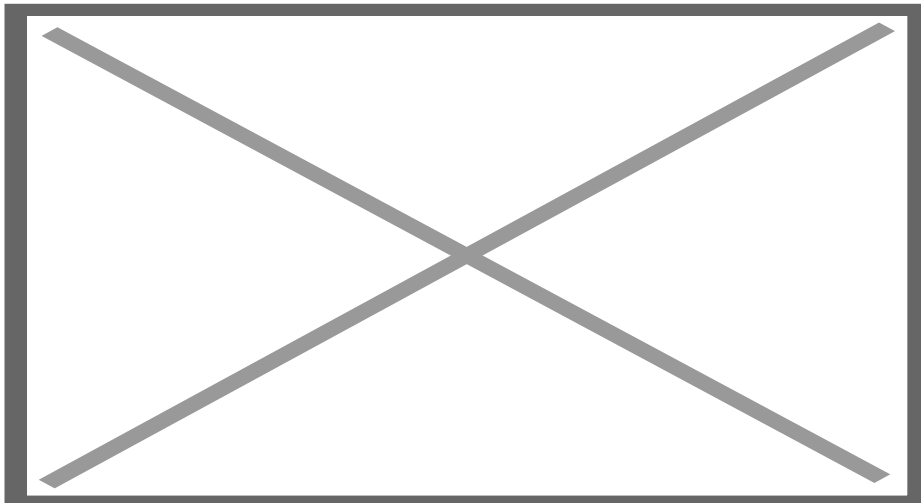
The DataScale program aims to provide students with dual expertise in:

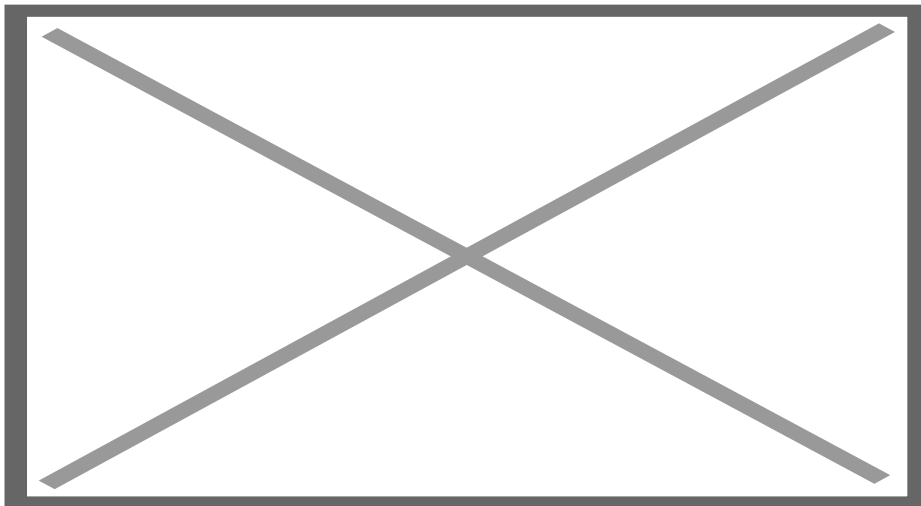
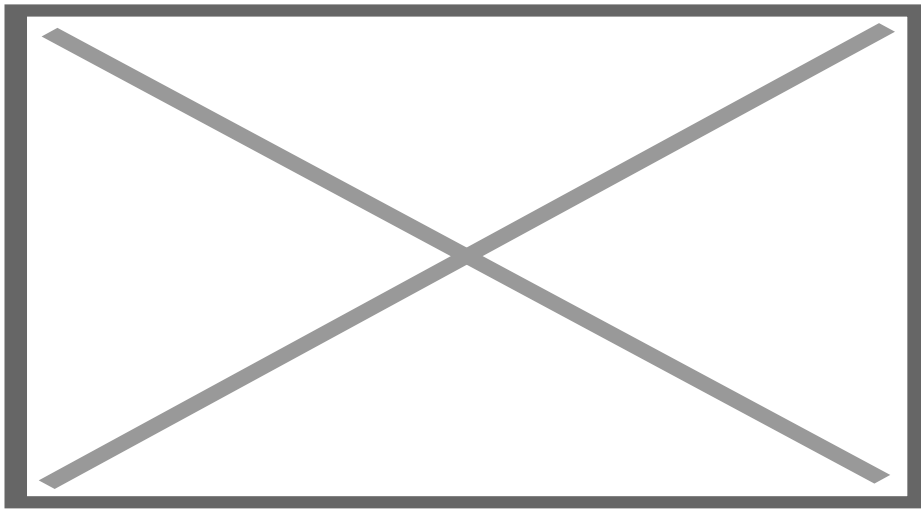
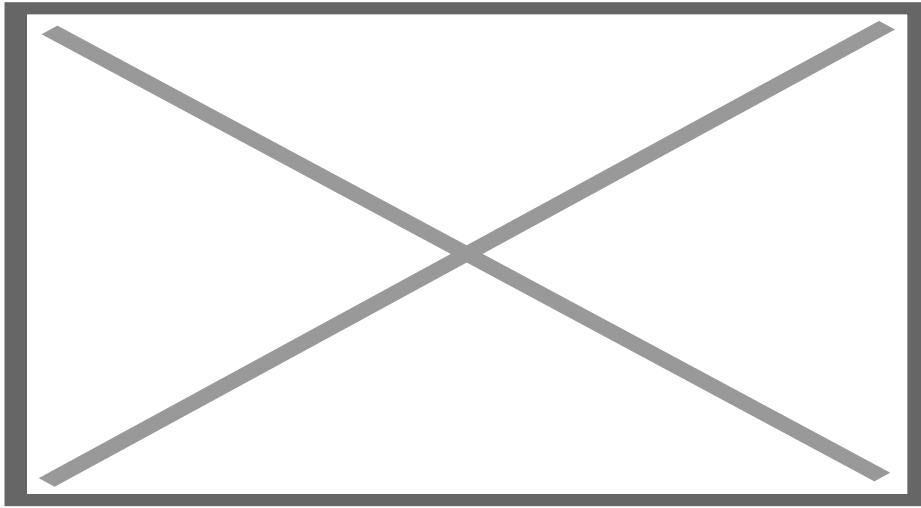
- **Data engineering:** design, administration, security, and optimization of multi-scale architectures (data centers, connected devices).
- **Data analysis:** integration and quality management of heterogeneous data, data mining, machine learning, and AI.

The program opens the way to a wide range of career opportunities (data scientist, data engineer, machine learning engineer, or AI analyst). Graduates may pursue doctoral studies in research laboratories or join service companies, startups, large corporations, or research structures.

[Télécharger la fiche détaillée](#)

[Apply Now](#)





■ ■

Information

- [Overview](#)
- [Application requirements](#)
- [Schedule](#)

Présentation

Skills

- Deploy, operate, and manage large-scale data infrastructures.
- Design and develop processing pipelines that transform raw data into knowledge.
- Extract, analyze, and leverage information and knowledge stored in massive data management infrastructures (big data).
- Master and deploy machine learning and AI techniques and algorithms.
- Develop and deploy service-oriented data management applications.
- Produce and present a state of the art, and carry out a scientific research approach.

Objectives

The DataScale program aims to provide students with dual expertise in:

- Data engineering: design, administration, security, and optimization of multi-scale architectures (data centers, connected devices).
- Data analysis: integration and quality management of heterogeneous data, data mining, machine learning, and AI.

The program opens the way to a wide range of career opportunities (data scientist, data engineer, machine learning engineer, or AI analyst). Graduates may pursue doctoral studies in research laboratories or join service companies, startups, large corporations, or research structures.

- Data engineer
- IA analyst
- Data analyst
- Data scientist
- Chief Data Officer
- Database Administrator (DBA)
- Data Security Administrator (DSA)
- Information System Architect
- Architect / Designer / developer of Big Data, Web and IoT applications

Calendar

Start of programme (indicative date)

14/09/2026

End of programme (indicative date)

31/08/2027

Admission

Capacity

Available Places

Target Audience and Entry Requirements

The DataScale program is offered exclusively as an initial training course. Prerequisites correspond to a Master's level in computer science, with a solid background in database management. The typical profiles of students joining DataScale program are: students who have obtained a Master's degree in IT in France, students from partner engineering schools in their last year enrolled in a dual program, students holding an engineering or computer science Master's degree obtained outside France.

Période(s) de candidature

Inception Platform

From 16/02/2026 to 10/07/2026

Supporting documents

Compulsory supporting documents

Copy diplomas.

Motivation letter.

All transcripts of the years / semesters validated since the high school diploma at the date of application.

Certificate of French (compulsory for non-French speakers).

Curriculum Vitae.

Additional supporting documents

Supporting documents :

- Residence permit stating the country of residence of the first country
- Or receipt of request stating the country of first asylum
- Or document from the UNHCR granting refugee status
- Or receipt of refugee status request delivered in France
- Or residence permit stating the refugee status delivered in France
- Or document stating subsidiary protection in France or abroad
- Or document stating temporary protection in France or abroad.

Programme

Tronc Commun

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	discovery/practical class
Machine Learning	2.5	Semestre 1	21						
■									

Machine Learning

ECTS : 2.5

Semester: Semestre 1

Detail

Lecture: 21

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Notions de statistiques et d'algèbre linéaire.

Type of assessment

Evaluation Continue non Intégrale

[Qualité des données](#)

2.5 Semestre 1 21

■

Qualité des données

ECTS : 2.5

Semester: Semestre 1

Detail

Lecture: 21

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Concepts fondamentaux de la gestion de données.

Type of assessment

Evaluation Continue non Intégrale

[Modèles et éco-systèmes](#)
[Post-Relationnels](#)

2.5 Semestre 1 21 0



Modèles et éco-systèmes Post-Relationnels

ECTS : 2.5
Semester: Semestre 1
Detail
Lecture: 21
Directed study: 0
Language(s) of instruction
Français
Remote teaching
non
Prerequisites

Maîtrise des concepts des bases de données relationnelles et des systèmes de gestion de bases de données, maîtrise d'un langage de programmation orienté-objet (Java, Python, ...), maîtrise des outils systèmes (shell, ...) et des outils de développement.

Type of assessment
Evaluation Continue non Intégrale

Intégration de données	2.5	Semestre 1	21
--	-----	------------	----



Intégration de données

ECTS : 2.5
Semester: Semestre 1
Detail
Lecture: 21
Language(s) of instruction
Français
Remote teaching
non
Prerequisites

Bases de données : création et manipulation de bases de données relationnelles, mécanismes internes des SGBD relationnels.

Type of assessment
Evaluation Continue non Intégrale

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	discovery/practical class
Architectures orientées Services	2.5	Semestre 1	21	0					
■									

Architectures orientées Services

ECTS : 2.5

Semester: Semestre 1

Detail

Lecture: 21

Directed study: 0

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Bonne connaissance des concepts de base de la programmation orientée-objet, et du développement d'applications Web.

Type of assessment

Evaluation Continue non Intégrale

[Architectures des gestionnaires de données](#)

2.5	Semestre 1	21	0	0
-----	------------	----	---	---

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	dis
----------	------	----------	---------	----------------	-----------------	------------------------	-------------------------	--------------------------------	-----

■

Architectures des gestionnaires de données

ECTS : 2.5

Semester: Semestre 1

Detail

Lecture: 21

Practical study : 0

Directed study: 0

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Niveau M1 informatique ou 4ème année d'école
d'ingénieur ayant reçu une formation en :

- bases de données : concepts
fondamentaux, SGBD relationnels,
principes de base de l'indexation et de
l'évaluation de requêtes
- systèmes distribués : concepts
fondamentaux, principes de base du
contrôle de concurrence et de la réplication

Type of assessment

Evaluation Continue non Intégrale

Data Analysis

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	dis
----------	------	----------	---------	----------------	-----------------	------------------------	-------------------------	--------------------------------	-----

[Fouille de
données](#)

2.5 Semestre 1 21



Fouille de données

ECTS : 2.5
Semester: Semestre 1
Detail
Lecture: 21
Language(s) of instruction
Français
Remote teaching
non
Prerequisites

Notions de statistiques et d'analyse de données.

Type of assessment
Evaluation Continue non Intégrale

Web sémantique	2.5	Semestre 1	21
--------------------------------	-----	------------	----



Web sémantique

ECTS : 2.5
Semester: Semestre 1
Detail
Lecture: 21
Language(s) of instruction
Français
Remote teaching
non
Prerequisites

Maîtrise des concepts de la gestion de données, maîtrise d'un langage de programmation (Java ou Python).

Type of assessment
Evaluation Continue non Intégrale

Modélisation de processus métiers	2.5	Semestre 1	21	0
---	-----	------------	----	---



Modélisation de processus métiers

ECTS : 2.5
Semester: Semestre 1
Detail
Lecture: 21
Directed study: 0
Language(s) of instruction
Français
Remote teaching
non
Prerequisites

Bonne connaissance des concepts de base des systèmes d’information, les architectures SOA, et du développement d’applications Web.

Type of assessment
Evaluation Continue non Intégrale

[Analyse de masses de données de mobilité](#)
2.5 Semestre 1 21



Analyse de masses de données de mobilité

ECTS : 2.5
Semester: Semestre 1
Detail
Lecture: 21
Language(s) of instruction
Français
Remote teaching
non
Prerequisites

Advance database techniques (query processing, data integration and quality) Machine learning (recommended)

Type of assessment
Evaluation Continue non Intégrale

Professionnalisation

Subjects	ECTS	Semester	Lecture study	directed practical class	Lecture/directed study	Lecture/practical class	directed study/practical class
Séminaires Industriels et Recherche	1	Semestre 2	20				
■							

Séminaires Industriels et Recherche

ECTS : 1

Semester: Semestre 2

Detail

Directed study: 20

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Aucun

Type of assessment

Evaluation Continue non Intégrale

Projet Programmation	1.5	Semestre 2	1.5
--------------------------------------	-----	------------	-----

■

Projet Programmation

ECTS : 1.5

Semester: Semestre 2

Detail

Directed study: 1.5

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

UE 'projet conception'

Type of assessment

Evaluation Continue non Intégrale

Projet Conception	1.5	Semestre 2	1.5
-----------------------------------	-----	------------	-----

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	practical class	directed study/practical class
■								

Projet Conception

ECTS : 1.5

Semester: Semestre 2

Detail

Directed study: 1.5

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Suivi des UE de tronc commun DataScale

Type of assessment

Evaluation Continue non Intégrale

Anglais	3	Semestre	27
		2	

■

Anglais

ECTS : 3

Semester: Semestre 2

Detail

Directed study: 27

Language(s) of instruction

Français

Remote teaching

non

Type of assessment

Evaluation Continue non Intégrale

Régulations des données et des contenus numériques	1	Semestre	9
		2	

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	distance learning courses
----------	------	----------	---------	----------------	-----------------	------------------------	-------------------------	--------------------------------	---------------------------

Régulations des données et des contenus numériques

ECTS : 1

Semester: Semestre 2

Detail

Lecture: 9

Language(s) of instruction

Français

Remote teaching

non

Type of assessment

Evaluation Continue non Intégrale

Stage

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	distance learning courses
----------	------	----------	---------	----------------	-----------------	------------------------	-------------------------	--------------------------------	---------------------------

[Stage](#) 22 Semestre 2 2

Stage

ECTS : 22

Semester: Semestre 2

Detail

Directed study: 2

Language(s) of instruction

Français

Remote teaching

non

Type of assessment

Evaluation Continue non Intégrale

Data Engineering

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	distance learning courses
----------	------	----------	---------	----------------	-----------------	------------------------	-------------------------	--------------------------------	---------------------------

[Sécurité des données corporate et personnelles](#) 2.5 Semestre 1 21

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	d
----------	------	----------	---------	----------------	-----------------	------------------------	-------------------------	--------------------------------	---

■

Sécurité des données corporate et personnelles

ECTS : 2.5

Semester: Semestre 1

Detail

Lecture: 21

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Bonnes connaissances des SGBD

Concepts de base de la cryptographie (un rappel sera effectué)

Type of assessment

Evaluation Continue non Intégrale

[Fondements](#)

[des langages](#)

[d'interrogation](#)

[de données](#)

2.5

Semestre

1

21

Fondements des langages d'interrogation de données

ECTS : 2.5

Semester: Semestre 1

Detail

Lecture: 21

Language(s) of instruction

Anglais

Remote teaching

non

Prerequisites

Notions de logique

Learning objectives

The course aims to provide a principled introduction to query processing, its logical foundations, and the recent query language evolution. It will cover the basics of graph processing: relational querying, Datalog for expanding its abilities to handle graphs, and Cypher, the most commonly currently used graph query language.

Overall organisation

It introduces key notions from first-order logic and gives a gentle overview of the fundamentals of the relational data model, with a focus on relational algebra and calculus. It then delves deeper into an analysis of the logical foundations of query languages. As such, it provides a detailed presentation of Datalog [6], which is a recursive query language that extends relational algebra with recursion, and that has been used to express a wide spectrum of modern data management tasks. We cover the standard Datalog language, its extension with negation, as well as application-specific dialects (with aggregates, custom predicates, and internalized transitive closure). Focusing on some of these fragments, we survey their various semantics [1, 8, 2]. We then discuss corresponding evaluation techniques and optimizations [5]. In the second part of the course, we concentrate on the connections between Datalog and graph query languages [7, 2]. We illustrate these through practical, interactive hands-on examples centered around Neo4j's Cypher query language. We will unravel basic and

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	d
Gestion de données et de services dans le cloud	2.5	Semestre 1	21	0					
■									

Gestion de données et de services dans le cloud

ECTS : 2.5

Semester: Semestre 1

Detail

Lecture: 21

Directed study: 0

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Bonne connaissance des architectures à bases de services, des concepts fondamentaux des bases de données, des applications client/serveur.

Type of assessment

Evaluation Continue non Intégrale

[Indexation et visualisation de données massives](#)

2.5 Semestre 1 21

Subjects	ECTS	Semester	Lecture	directed study	practical class	Lecture/directed study	Lecture/practical class	directed study/practical class	d
----------	------	----------	---------	----------------	-----------------	------------------------	-------------------------	--------------------------------	---

■

Indexation et visualisation de données massives

ECTS : 2.5

Semester: Semestre 1

Detail

Lecture: 21

Language(s) of instruction

Français

Remote teaching

non

Prerequisites

Bonne connaissance des SGBD et des éco-systèmes post-relationnels.

Learning objectives

La formation présente un état de l'art des concepts d'indexation et de visualisation de données massives. Elle intègre également la mise en oeuvre de la pile Elastic Stack (Elasticsearch, Logstash, Kibana) afin de permettre aux experts métiers de questionner les différents fichiers de traces produits par un système d'information. Il est alors possible de gérer, d'exploiter et de visualiser les logs de tous type d'application et d'OS.

Overall organisation

- Connaître les principes fondamentaux de l'indexation et de la visualisation de données massives
- Comprendre la pile ELK (Elasticsearch, logstash, Kibana) et quand l'utiliser
- Installer et configurer Elasticsearch, logstash et Kibana
- Savoir utiliser Elasticsearch, logstash et Kibana pour indexer, chercher et visualiser des données et des documents (projet)
- Produire des tableaux de bord
- Intégrer la pile avec Hadoop et Spark et Kafka

Type of assessment

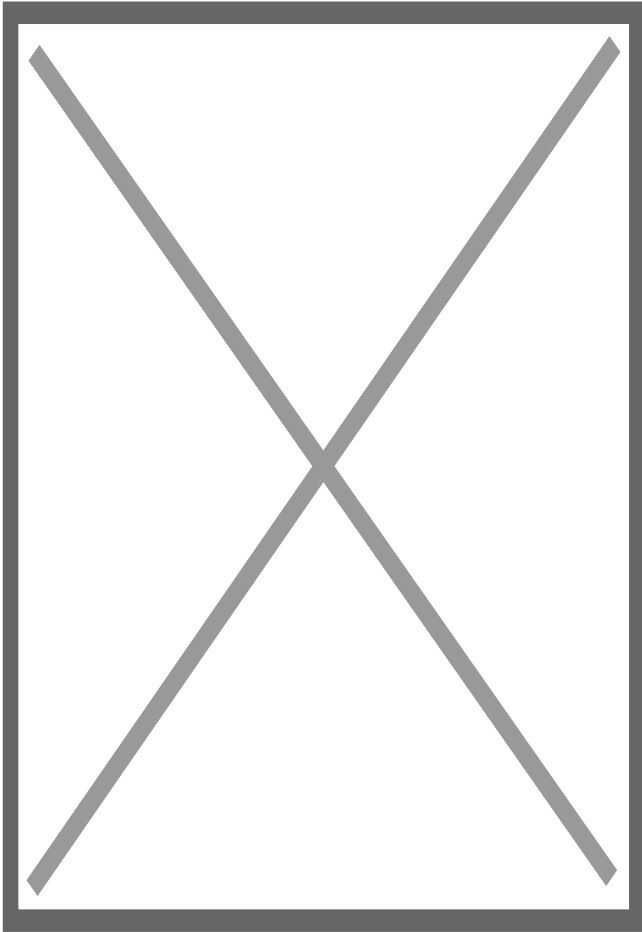
Evaluation Continue non Intégrale

Teaching Location(s)

VERSAILLES

Training campus

Javascript est requis pour visualiser cette carte.



Versailles

Versailles

Bus 8 - 10 - 13 - 105 - 108 - 471

Versailles Rive Droite - Ligne L

RER C Versailles Château Rive gauche

[Voir](#)

Student restaurant (CROUS)

Library

Student residence

Contact

- **Zoubida KEDAD**

PR
- **Mustapha LEBBAH**

PR
- **Clémence Perez**

Secrétaire Département d'informatique

<http://www.universite-paris-saclay.fr/formation/master/informatique/m2-datascale-gestion-de-donnees-et->

