Libera Università Internazionale
degli Studi Sociali Guido Carli

LUISS

Syllabus

Academic Year	2026/2027
Program	Data Science and Management
Course	Computing Infrastructures for Data Science
Term	I semester
Year	1
SSD	INFO-01/A
Credits	4

INSTRUCTIONAL GOALS	The course introduces the fundamental concepts of computational infrastructures for data science. It aims to provide students with a solid understanding of hardware and software architectures, distributed systems, cloud computing, and tools for efficient management of large-scale data. Particular attention is devoted to scalability, reliability, performance, and resource optimization, enabling students to design and use modern data-intensive computing infrastructures.
INTENDED LEARNING OUTCOMES They describe what a learner is expected to know, understand and be able to demonstrate after completion of a learning path.	**Knowledge and understanding**: Students will acquire knowledge of the main computational infrastructures used in data science, including computing architectures, storage systems, distributed and parallel systems, cloud platforms, and data processing frameworks. **Applying knowledge and understanding**: Students will be able to select and use appropriate computational infrastructures and tools for data science tasks, configure basic distributed and cloud environments, and manage data processing workflows efficiently. **Making judgements**: Students will develop the ability to critically evaluate different infrastructure solutions, assessing trade-offs in terms of performance, scalability, cost, and reliability for specific data science applications. **Communication skills**: Students will be able to clearly communicate technical concepts related to computing infrastructures, using appropriate terminology, and to interact effectively with both technical and non-technical stakeholders. **Learning skills**: The knowledge acquired during the course will enable students to independently update their skills and adapt to evolving technologies in computing infrastructures and data science ecosystems.
Pre-requisites	Basic knowledge of programming and introductory concepts of computer science. Familiarity with data analysis is recommended.

Course content	• Introduction to computing infrastructures for data science • Computer architectures: CPUs, GPUs, accelerators, and memory hierarchies • Storage systems and data management • Distributed and parallel computing models • Big data processing frameworks (e.g., Hadoop, Spark) • Cloud computing: service models, architectures, and platforms • Containerization and orchestration (e.g., Docker, Kubernetes) • Scalability, reliability, and fault tolerance • Performance evaluation and optimization
Reference Books	• T. White (2015), Hadoop: The Definitive Guide (O'Reilly Media) • M. Zaharia et al. (2016), Spark: The Definitive Guide (O'Reilly Media) • Additional slides, scientific articles, and teaching materials provided during the course (on MyLuiss).
Teaching Methods	• Frontal lectures • Hands-on labs and practical exercises • Case studies and project-based activities
Assessment	• Practical final project / assignments (1/3) • Mid-term written or practical exam (1/3) • Oral or final written exam (1/3)