ENGINEERING FOR SAFETY OF CRITICAL INDUSTRIAL AND CIVIL

(Lecce - Università degli Studi)

Teaching DATA ENGINEERING AND SECURITY FOR CYBER PHYSICAL SYSTEMS

GenCod A007225

Owner professor ANTONELLA LONGO

Teaching in italian DATA ENGINEERING AND SECURITY FOR CYBER PHYSICAL

Teaching DATA ENGINEERING AND SECURITY FOR CYBER PHYSICAL

SSD code ING-INF/05

Reference course ENGINEERING FOR SAFETY OF CRITICAL INDUSTRIAL AND

Course type Laurea Magistrale

Credits 9.0

Teaching hours Front activity hours: 81.0

For enrolled in 2024/2025

Taught in 2024/2025

Course year 1

Language ENGLISH

Curriculum INDUSTRIAL ENGINEERING SYSTEMS

Location Lecce

Semester First Semester

Exam type Oral

Assessment Final grade

Course timetable https://easyroom.unisalento.it/Orario

BRIEF COURSE DESCRIPTION

Data engineering is becoming a foundational skill, leveraged by the incremental growth of cyber - physical systems. Critical Infrastructures (i.e. those infrastructures which deliver essential services in the society, like energy infrastructures, mobility, water, data transmission, health, etc.) are more and more based on the use of digital technologies, which sense, monitor and forecast physical behaviours and/or the impact of some risks. The widespread definition of Digital Twins is currently based on the concept of real time exchange of data between the physical asset and the corresponding digital shadow.

The aim of this course is to provide students with the basic skills on how physical systems produce data, which can be collected, transformed and analyze and how complex organizations like critical infrastructures providers govern data.

Starting from the basic concepts related to the digital technologies for the automation of physical infrastructures, the course will provide a complete overview of the data engineering lifecycle. All the aspects related to data analysis will be faced, including data generation, ingestion, orchestration, transformation, storage, and governance. These aspects are critical in any data environment regardless of the underlying technology.

Essentially, the goals of the course are the following:

1. Provide the foundational knowledge of digital technologies for automating physical systems
2. Get a concise overview of the entire data engineering landscape, starting with the generation and provisioning of data from the real world to the data analysis and visualization.
3. Assess data engineering problems, in critical infrastructures' scenarios
4. Use the data engineering lifecycle to design and build pipelines of data processing
5. Incorporate data governance and security across the data engineering lifecycle

REQUIREMENTS

No previous skills are required. The use of computer and the tools of office automation are

COURSE AIMS

Knowledge and understanding. Students must have a solid background related to the basics of data engineering in cyber-physical systems:

- They must have the basis to think analytically, creativelly and critically and being able to create abstraction and problem solving skills to cope with complex systems
- They must have a basic knowledge of design and implementation of data management systems
- They must have the tools to design transactional and analytical databases applied to different contexts
- They must have the skills to argument data in different scenario, the tools for managing them, together with its impact.

Applying knowledge and understanding. After the course the student will be able to:

- Describe the model and frameworks of a data engineering; illustrate the main components of data lifecycle from the technical and application perspective and the impact on the critical infrastructures management.
- Distinguish conceptual, logical and physical models in data management.
- Model Online Transaction processing systems from a data perspective, distinguishing among ER models, relational models and physical models
- Model Online Analytical processing systems form a data perspective, distinguishing among DFM, Snowflakes and physical models, being able to describe the relationships among them and the processes
- Explore data sets, applying proper techniques, including the integration of different data sources

Making judgements. Students are guided to critically approach the topics treated during the class, to compare different solutions to a problem, to identify and propose the most effective or efficient solution in an automous way.

Communication. Students will learn to communicate wiith heterougenoeus audiences, explaining their position, in logical, coherent and effecitve way. During the course students will be provided with domain specific vocabulary and the proper scientific knowledge and methods to expose and argument in precise and formal way the main topics related to data management and information system

Learning skills. Students must acquire the critical ability to autonomously relate to the typical problems of data management and, in general, cultural issues related to data governance and management. They should be able to develop an approach to independently structure knowledge and methods learnt with a view to possible follow up of studies at higher (doctoral) level or in the broader perspective of cultural and professional self-improvement of lifelong learning. Therefore, students should be able to switch their learning approach according to different learning sources and the objectives they must achieve in terms of results and audience

TEACHING METHODOLOGY

The course aims to provide students with tools and knowledge for data engineeirng applied to critical infrastructures. The course consists of frontal lessons and classroom hands on exercises. The frontal lessons are aimed at improving students' knowledge and understanding through the presentation of theories, models and methods; students are invited to participate in the lesson with autonomy of judgement, by asking questions and presenting examples. The exercises are aimed at using tools which supports the models and approaches presented

ASSESSMENT TYPE	The exam is an interview made up of both practical and descriptive aspects. The practical part aims at evaluating to what extent the student has: 1) the ability to design data models according to the methodologies presented during the call, 2) reasoning about his/her choices and the capacity to integrate different concepts and tools. The descriptive part follows the practical part and is aimed to verify to what extent the student has gained knowledge and understanding of selected topics and he is able to communicate them. Optionally students can develop specific use cases, agreed with the teacher

FULL SYLLABUS

1. Introduction to digital systems: the Von Neumann model
2. Definition, design and development of a cyber- physical system.
3. Data format and models: conceptual models, logical models, physical models
4. Relational and non relational databases
5. Data Quality in real world use cases
6. Data Governance

REFERENCE TEXT BOOKS

1. “Fundamentals of Database Systems", 6th Edition, Elmasri, Navathe, Addison-Wesley
2. “Datawarehouse Design - Modern Principles and Methodologies”, Matteo Golfarelli, Stefano Rizzi, McGrawHill
3. “Information Systems” Paige Baltzan, 4th edition, Mc Graw Hill Education
4. Teaching materials provided at the course