Fakultät für Ingenieurwesen
Facoltà di Ingegneria
Faculty of Engineering

COURSE DESCRIPTION – ACADEMIC YEAR 2024/2025

Course title	**Information Retrieval**
Course code	76057
Scientific sector	INF/01
Degree	Master in Computational Data Science (LM-18)
Semester	1
Year	2
Credits	6
Modular	No

Total lecturing hours	40
Total lab hours	20
Attendance	The attendance is not compulsory, but students are highly encouraged to attend
Prerequisites	Programming and algorithm data structures skills, Linear algebra, probability theory, basic machine learning concepts.
Course page	Teams and https://ole.unibz.it/

Specific educational objectives	The course belongs to the type "caratterizzanti – discipline informatiche" in the study path without curriculum". The objective of this course is to present the scientific underpinnings of the field of Information Retrieval (IR). The student will study fundamental, mathematically sophisticated IR concepts first and then more advanced techniques for information filtering and decision support, including transformer-based solutions and LLMs. This course provides students with a rich and comprehensive catalogue of information search and text processing techniques that can be exploited for the design and implementation of modern IR applications.

Lecturer	Andrea Rosani
Contact	Via Bruno Buozzi 1, Room B1.5.24, andrea.rosani@unibz.it
Scientific sector of lecturer	INF/01
Teaching language	English
Office hours	During the lecture time span, arrange beforehand by email.
Lecturing Assistant (if any)	-
Contact LA	-
Office hours LA	-
List of topics	• Web and mobile search • Boolean and vector-space retrieval models • Efficient document indexing, document mining and topic modelling • Traditional and machine learning-based ranking approaches • Foundation models • Evaluation of Information Retrieval Systems

Teaching format	Frontal lectures, exercises, lab, seminars.

Learning outcomes	Knowledge and understanding: • D1.4 - Sound basic knowledge of storing, querying and managing large amounts of data and the associated languages, tools and system. Applying knowledge and understanding: • D2.2 - Ability to address and solve a problem using scientific Methods Making judgments • D3.2 - Ability to autonomously select the documentation (in the form of books, web, magazines, etc.) needed to keep up to date in a given sector. Communication skills • D4.1 - Ability to use English at an advanced level with particular reference to disciplinary terminology.

Assessment	Final Project + oral exam The project will consist of the design of an IR system in a specific application domain selected by the students. The project domain, the attacked problem, the techniques, and the obtained results must be described in a report (max. 10 pages). The project can be done in groups of 2-3 people. The oral exam is composed by the discussion of the project and some individual questions on the content of the course.
Assessment language	English
Assessment Typology	Monocratic
Evaluation criteria and criteria for awarding marks	Evaluation criteria - Project: 50% of the mark - Final oral exam: 50% of the mark. Important note: both project and exam are required to be passed. **Criteria for awarding marks** Oral exam: ability to present and explain information retrieval concepts, methods and algorithms. ability to select appropriate solutions for IR problems. Project: ability to implement data workflow to apply IR to real-world problems, correctness and clarity of the solution, experimental results, ability to solve IR problems with the appropriate technique.

Required readings	The suggested book for the introduction to information retrieval topics is:

	• C. D. Manning, P. Raghavan and H. Schutze. Introduction to Information Retrieval, Cambridge University Press, 2008. (Online: http://informationretrieval.org) Papers about the most recent advancements with regards to algorithms, information access modalities and interfaces will be provided during the course in electronic format. Copy of the slides will be available as well. Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it
Supplementary readings	Scientific papers and additional books will be specified on each course topic during lectures. Rao, Delip, and Brian McMahan. Natural language processing with PyTorch: build intelligent language applications using deep learning. " O'Reilly Media, Inc.", 2019. Tunstall, Lewis, Leandro Von Werra, and Thomas Wolf. Natural language processing with transformers. " O'Reilly Media, Inc.", 2022.
Software used	Python as programming language.