

COURSE DESCRIPTION – ACADEMIC YEAR 2023/2024

Course title	Information Retrieval
Course code	73002
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	Attending lectures is not compulsory. Final marks will be awarded on the basis of the outcome of the final project and assignment (a paper to study and present) that will be given during the lecture period. Those who are not attending the lectures must contact the lecturer to receive their mandatory assignments.
Prerequisites	Programming and algorithm data structures skills, Linear algebra, probability theory, basic machine learning concepts (example, attribute, instance-based learning, cross validation and hyperparameter selection, evaluating predictions). Good to have knowledge of basic HCI and database principles.
Course page	Teams and https://ole.unibz.it/
Specific educational objectives	<p>The course belongs to the type "caratterizzanti – discipline informatiche" in the study path without curriculum".</p> <p>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. During the course lectures and seminars, the students will acquire and develop skills in topics that are timely and demanded by leading companies in the industry sector. A selection of these topics is here given: personalization of retrieval results; interactive information retrieval; and conversational IR models and interfaces.</p> <p>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.</p>
Lecturer	Mauro Dragoni
Contact	Faculty of Engineering, piazza Domenicani, 3 Bolzano, Office 1.04, Mauro.Dragoni@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	<ul style="list-style-type: none"> • Efficient document indexing • Boolean and vector-space retrieval models • Web and mobile search • Document mining and topic modelling • Traditional and machine learning-based ranking approaches • Evaluation of Information Retrieval Systems
Teaching format	Frontal lectures, exercises, lab, seminars.
Learning outcomes	<p>Knowledge and understanding:</p> <ul style="list-style-type: none"> • D1.4 - Sound basic knowledge of storing, querying and managing large amounts of data and the associated languages, tools and system. <p>Applying knowledge and understanding:</p> <ul style="list-style-type: none"> • D2.2 - Ability to address and solve a problem using scientific Methods <p>Making judgments</p> <ul style="list-style-type: none"> • 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. <p>Communication skills</p> <ul style="list-style-type: none"> • D4.1 - Ability to use English at an advanced level with particular reference to disciplinary terminology.
Assessment	<p>Final Project (individual), 65% of mark Seminars, 35% of mark</p> <p>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 (formatted according to the ACM SIGCHI single column template, max. 5000 words). System prototype implementation is mandatory. The project results will be presented at the end of the course. The project is aimed at assessing to what extent the student has achieved the previously presented learning outcomes (applying knowledge and understanding, making judgments, communication skills).</p> <p>The student will be required, during the lecture period, to give one seminar to present two scientific IR papers. The seminar is aimed at assessing to what extent the student has achieved the learning outcomes related to: applying knowledge and understanding, making judgments, communication and learning skills.</p> <p>The exam consists in presenting the project in a complete way by describing both the implementation and the theoretical background leading to the implementation decisions. It will assess to what extent the student has achieved the above-mentioned learning outcomes related to: knowledge and understanding, applying knowledge and understanding, and learning skills.</p>

Assessment language	English
Assessment Typology	Monocratic
Evaluation criteria and criteria for awarding marks	<p>The seminar will be evaluated at the end of the semester, and it is a prerequisite for attending the written exam. The student must pass each single component of the exam (project and seminar), by obtaining at least 50% of the mark.</p> <p>PROJECT EVALUATION CRITERIA: clarity of the presentation, novelty of the problem and the solution, appropriateness of the methods, significance of the solution and capability to refer to the methods and techniques discussed in the course.</p> <p>SEMINAR EVALUATION: clarity of the presentation and capability to reply to clarification questions of the teacher.</p>
Required readings	<p>The suggested book for the introduction to information retrieval topics is:</p> <ul style="list-style-type: none"> • C. D. Manning, P. Raghavan and H. Schütze. Introduction to Information Retrieval, Cambridge University Press, 2008. (Online: http://informationretrieval.org) <p>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.</p> <p>Subject Librarian: David Gebhardi, David.Gebhardi@unibz.it</p>
Supplementary readings	Scientific papers and additional books will be specified on each course topic during lectures.
Software used	Java and Python as programming languages. Solr, Lucene, and RASA as software tools.