Course guide
270414 - IBD - Introduction to Databases

Last modified: 10/07/2024

Unit in charge: Barcelona School of Informatics
Teaching unit: 747 - ESSI - Department of Service and Information System Engineering.

Degree: BACHELOR'S DEGREE IN ARTIFICIAL INTELLIGENCE (Syllabus 2021). (Compulsory subject).

Academic year: 2024 **ECTS Credits:** 6.0 **Languages:** Catalan

LECTURER

Coordinating lecturer: PETAR JOVANOVIC

Others: Primer quadrimestre:
PETAR JOVANOVIC - 11, 12
GERARD PONS RECASENS - 11
ANNA QUERALT CALAFAT - 12

PRIOR SKILLS

To know the data structures in internal memory.
To be able to implement programs of medium complexity.

DEGREE COMPETENCES TO WHICH THE SUBJECT CONTRIBUTES

Specific:
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.

Generical:
CG1. To ideate, draft, organize, plan and develop projects in the field of artificial intelligence.
CG2. To use the fundamental knowledge and solid work methodologies acquired during the studies to adapt to the new technological scenarios of the future.

Transversal:
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.

Basic:
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.

TEACHING METHODOLOGY

Theory classes / problems
Autonomous learning: To prepare classes the student may have to read and understand materials and / or notes indicated by the teacher. Afterwards in class, the student needs to review and solve exercises on the topic of study.
Theory classes In lectures the teachers present a part of the contents of the subject. Normally, teachers use transparencies that students would be advised to obtain before classes, in order to do a better follow-up.
Problems classes In problem classes, students solve exercises about content presented during theory classes. These exercises are done in teams of two students according to a cooperative learning technique.
Evaluation. In four of the problem classes, students will solve an exercise that will be collected and evaluated by the teacher.

Laboratory classes
Autonomous learning: The contents that are worked on in the laboratory classes will be studied autonomously by the students. Each week before in the laboratory class students will have a homework assignment that will end with the resolution of a moodle / LearnSQL quiz.
Laboratory classes: Class work will be in teams of 2 students. Students have the opportunity to share doubts with their teammate about the work they have done at home, and if necessary, to ask questions that are not resolved to the teacher. Next the students do the activities that the teacher has indicated and finally solve the class questionnaire.
Assessment: In each class, the students answer a question individually to evaluate the work done at home previously and during the class. The evaluation is also based on the exercices solved during the class.

LEARNING OBJECTIVES OF THE SUBJECT

1.To know the objectives of a database management system and their architecture.
2.To understand the database relational model, their languages (SQL and relational algebra) and the usual components of a relational database.
3.To be able to define, create and manipulate usual relational database components.
4.To be able to build programs to manage relational databases.
5.To know the different available formats for semistructured data, and know how to write SQL queries over them.
6.To be able to apply some defined quality criteria to choose between several SQL statements, database components, or programs, that manage a database and implement the same functionality.
7.To be able to apply some defined quality criteria to choose between several SQL statements, database components, or programs, that manage a database and implement the same functionality.
8.To have a general vision of how the design of a database should be included in a software development process.
9.To be able to obtain a database relational model starting from a conceptual models in UML.
10.To know the concept of database transaction and its implications.
11.To know how to identify the different types of interference that can occur between database transactions and their relationship with the isolation levels that defines the SQL
12.To know the locking concurrency control technique.
13.To know the possible physical structures for storing data and its implications for in terms of efficiency.
14.To know the access methods to data and its implications in terms of efficiency.
16.To have a general vision of what a database is, what is a database model, the types of users of databases and which are the categories of databases languages.

STUDY LOAD

Type	Hours	Percentage
Self study	90,0	60.00
Hours small group	30,0	20.00
Hours large group	30,0	20.00

Total learning time: 150 h

CONTENTS

Introduction

Description:
Database concept. Database design and models. Types of users. Categories of languages. Concept of database management system (DBMS). Desirable goals for databases that DBMSs must provide. Architecture of the DBMS.

The relational model

Description:
Objectives and origin. Structure of data with which the relational databases are built. Operations provided by the relational model to manipulate and query the data. Integrity rules to be met by the data in a relational database.

Languages: Relational algebra and SQL

Description:
Introduction. Relational algebra: operations of relational algebra; queries. SQL: table creation; insertion, deletion and modification of rows in a table; queries on a database. Considerations about the implementation of queries.

Logical database components

Description:
Concept of a logical database component: data and control components. Introduction to the data components: schemes, tables and domains, assertions and views. Introduction to the control components: stored procedures, triggers and privileges.

Semistructured data formats and SQL extensions to query them

Description:
Introduction to the different semi-structured data formats. SQL extensions to query semi-structured data.

SQL Programming

Description:
Programming in Python and DataFrames. Considerations and quality criteria in the design and implementation of programs that access databases.

Transactions and concurrency

Description:
Concept of transaction. ACID properties of transactions. Interference between transactions. Serialitzability. Recoverability. Concurrency control techniques. Isolation Levels. Locking and isolation levels.

Introduction to the design of relational databases

Description:
Stages in the design of a database. Introduction to the understanding of simple UML conceptual models. Translation of simple UML conceptual models to relational model databases

Physical storage structures, access methods and optimization

Description:
Introduction. Access methods to perform queries and updates in a database. Costs of the different access methods. Introduction to Ouery Optimaztion

ACTIVITIES

Study of the database introduction

Specific objectives:
1, 16

Related competencies :
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.

Full-or-part-time: 8h
Self study: 4h
Theory classes: 2h
Laboratory classes: 2h

Study of the databases introduction

Specific objectives:
2

Related competencies :
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.

Full-or-part-time: 2h
Self study: 2h

Study of the data logical components

Specific objectives:
2, 3, 6

Related competencies :
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.

Full-or-part-time: 10h
Self study: 6h
Theory classes: 4h

Study of the introduction to design of relational databases

Specific objectives:
8, 9

Related competencies :
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.

Full-or-part-time: 8h
Self study: 4h
Theory classes: 4h

Study of transactions and concurrency

Specific objectives:
10, 11, 12

Related competencies :
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.

Full-or-part-time: 10h
Self study: 6h
Theory classes: 4h

Study of storage, access methods and optimization

Specific objectives:
13, 14

Related competencies :
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CG1. To ideate, draft, organize, plan and develop projects in the field of artificial intelligence.
CG2. To use the fundamental knowledge and solid work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.

Full-or-part-time: 17h
Self study: 7h
Theory classes: 10h

Study of the Relational Algebra and SQL

Specific objectives:
2, 3, 6

Related competencies :
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.

Full-or-part-time: 24h
Self study: 12h
Laboratory classes: 12h

Study of semistructured data models and SQL extensions to query them

Specific objectives:
5, 6, 7

Related competencies :
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CG1. To ideate, draft, organize, plan and develop projects in the field of artificial intelligence.
CG2. To use the fundamental knowledge and solid work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.

Full-or-part-time: 8h
Self study: 4h
Theory classes: 2h
Laboratory classes: 2h

Study of stored procedures and triggers

Specific objectives:
2, 3, 6

Related competencies :
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.

Full-or-part-time: 14h
Self study: 6h
Laboratory classes: 8h

Programming with SQL - Python and DataFrames

Specific objectives:
3, 4, 6

Related competencies :
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.

Full-or-part-time: 10h
Self study: 4h
Laboratory classes: 6h

Final exam

Specific objectives:
1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16

Related competencies :
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CG1. To ideate, draft, organize, plan and develop projects in the field of artificial intelligence.
CG2. To use the fundamental knowledge and solid work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.

Full-or-part-time: 20h
Self study: 18h
Guided activities: 2h

Reviews and resolution of doubts about the exams

Full-or-part-time: 3h
Self study: 3h

Midterm exam

Specific objectives:
1, 2, 3, 5, 9, 16

Related competencies :
CB3. That students have the ability to gather and interpret relevant data (usually within their area of ??study) to make judgments that include a reflection on relevant social, scientific or ethical issues.
CB2. That the students know how to apply their knowledge to their work or vocation in a professional way and possess the skills that are usually demonstrated through the elaboration and defense of arguments and problem solving within their area of ??study.
CE04. To design and use efficiently the most appropriate data types and structures to solve a problem.
CE08. To detect the characteristics, functionalities and components of data managers, which allow the adequate use of them in information flows, and the design, analysis and implementation of applications based on them.
CG1. To ideate, draft, organize, plan and develop projects in the field of artificial intelligence.
CG2. To use the fundamental knowledge and solid work methodologies acquired during the studies to adapt to the new technological scenarios of the future.
CT6. Autonomous Learning. Detect deficiencies in one's own knowledge and overcome them through critical reflection and the choice of the best action to extend this knowledge.
CT4. Teamwork. Be able to work as a member of an interdisciplinary team, either as a member or conducting management tasks, with the aim of contributing to develop projects with pragmatism and a sense of responsibility, taking commitments taking into account available resources.

Full-or-part-time: 16h
Self study: 14h
Guided activities: 2h

GRADING SYSTEM

The grade of the course is based on technical competencies:

- NPR: Problems grade. It is the average of the grades of the four problems exam.

- NLB: Active participation in laboratory sessions. The classes in which students have participated will be taken into account in case of successfully submission of the exercises proposed in the class. The grade will be calculated in proportion to the classes in which the students have actively participated.

- NEP - Partial exam grade.

- NEF: Final exam grade.

Final grade = Maximum (
NLB*0.2+NEP*0.30+NEF*0.35+NPR*0.15,
NLB*0,2+NEP*0.35+NEF*0,45
)

- For students who can take the re-assessment, the re-assessment exam mark will replace NEF and NEP. In any case, the final mark will be the maximum between the ordinary mark and the re-evaluation mark.

BIBLIOGRAPHY

Basic:
- Garcia-Molina, Hector; Ullman, Jeffrey D; Widom, Jennifer. Database systems : the complete book [on line]. Second edition, Pearson new international edition. Harlow, Essex: Pearson Education Limited, [2014] [Consultation: 14/03/2025]. Available on: https://ebookcentral-proquest-com.recursos.biblioteca.upc.edu/lib/upcatalunya-ebooks/detail.action?pq-origsite=primo&docID=5174436. ISBN 9781292024479.
- Ramakrishnan, Raghu; Gehrke, Johannes. Database management systems. 3rd ed., international ed. Boston [etc.]: McGraw-Hill, cop. 2003. ISBN 0071151109.
- Badia, Antonio. SQL for data science : data cleaning, wrangling and analytics with relational databases. Cham: Springer, 2020. ISBN 9783030575915.

Complementary:
- Gulutzan, Peter; Pelzer, Trudy. SQL-99 complete, really. Lawrence: R & D books, cop.1999. ISBN 0879305681.
- Liu, Ling; Özsu, M. Tamer. Encyclopedia of database systems. New York ; London: Springer, 2009. ISBN 9780387399409.

RESOURCES

Hyperlink:
- https://learnsql2.fib.upc.edu/moodle/course/view.php?id=83