

Data Governance

Code: 44749
ECTS Credits: 9

2025/2026

Degree	Type	Year
Archival Studies and Information Governance	OB	2

Contact

Name: Jordi Serra Serra

Email: jordi.serra.serra@uab.cat

Teachers

Marc Soriano Rue

Jordi Serra Serra

Teaching groups languages

You can view this information at the [end](#) of this document.

Prerequisites

1. It is recommended that the subject "Information Systems and Systems Architecture" has been previously taken
2. It is recommended that the subject "Information Description and Retrieval" has been previously taken

Objectives and Contextualisation

1. Know the data life cycle and its management.
2. Understand the context of data production.
3. Apply archival principles to data management.
4. Know and understand the main tools and systems for data management.
5. Know data management systems and databases.
6. Know data governance models, rules, and standards.
7. Know and understand the basic systems for data use, exploitation, and visualization.

Learning Outcomes

1. CA21 (Competence) Establish an organisation's quality data.
2. CA22 (Competence) Design the criteria and formats for managing the life cycle of an organisation's data.

3. KA30 (Knowledge) Describe the data life cycle.
4. KA31 (Knowledge) Identify types and formats of data.
5. KA32 (Knowledge) Recognise data governance systems: repositories, data architecture platforms, and database systems.
6. SA23 (Skill) Use the main data management instruments and systems.
7. SA24 (Skill) Apply data governance techniques in organisations.
8. SA25 (Skill) Apply archival principles to data management.

Content

- 1.1. Data in organizations (introduction)
- 1.2. Where are data produced?
 - 1.2.1. Ways of capturing and generating data (procedures, sensors, etc.)
 - 1.2.2. Models for structuring data (master, referential, etc.)
 - 1.2.3. Architectures for storage (types of databases)
- 1.3. How are data used?
 - 1.3.1. Data preparation
 - 1.3.1.1. Data formats to be cleaned
 - 1.3.1.2. Data cleansing
 - 1.3.1.3. Preparation for exploitation
 - 1.3.2. Exploitation and use of data
 - 1.3.2.1. Data visualization
 - 1.3.2.2. Advanced statistical analytics, or based on ML and AI
 - 1.3.2.3. Practical application of advanced analytics algorithms
 - 1.4. Integrated data governance
 - 1.4.1. Data identification and cataloging
 - 1.4.2. Data lineage control
 - 1.4.3. Virtualization of data access
 - 1.4.4. Legal and security aspects
 - 1.4.5. Links with archival science

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Theoretical sessions	45	1.8	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21
Type: Supervised			
Exercise 1: cleaning, debugging and preparation of a dataset.	30	1.2	CA21, CA22, CA21
Exercise 2: Creating a data visualization.	30	1.2	CA21, KA31, SA23, CA21
Exercise 3: running an advanced analysis on a data set.	20	0.8	CA21, SA23, SA24, CA21
Type: Autonomous			
Final test: test of general knowledge of the subject.	10	0.4	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21
Reading Materials	90	3.6	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25, CA21

The autonomous learning activities will be reading materials and preparing for the final general knowledge test of the course.

The directed activities will be theoretical lecture sessions.

The supervised activities will be 3 practical exercises to be done at home with the explanations received in class.

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Assessment

Continuous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Exercise 1: cleaning, debugging and preparing a dataset.	25% of the final grade	0	0	CA21, CA22
Exercise 2: creating a data visualization.	25% of the final grade	0	0	CA21, KA31, SA23
Exercise 3: execution of an advanced analysis on a set of data.	20% of the final grade	0	0	CA21, SA23, SA24
Final test: test of general knowledge of the subject.	30% of the final grade	0	0	CA21, CA22, KA30, KA31, KA32, SA23, SA24, SA25

Both exercises 1 and 2 will be worth 25% of the final grade. The 3rd will be worth 20% and the final exam 30%.

For this subject, the use of Artificial Intelligence (AI) technologies is allowed exclusively in support tasks, such as bibliographic or information search, text correction or translations, or the automatic generation of fictitious data sets for practices. AI may NOT be used, unless explicitly indicated by the teacher, for data cleaning, visualization generation, or code generation for advanced data analysis. Even in cases where its use has been indicated by the teacher, the student must clearly identify which parts have been generated with this technology, specify the tools used and include a critical reflection on how these have influenced the process and the final result of the activity. The lack of transparency in the use of AI in this assessable activity will be considered a lack of academic honesty and may lead to a partial or total penalty in the grade of the activity, or greater sanctions in serious cases.

Bibliography

Earley, S., & Henderson, D. (Ed.). (2017). DAMA-DMBOK: Data management body of knowledge (2nd edition). Data Management Association.

Ghavami, P. (2020). Big data management: Data governance principles for big data analytics (1a ed.). De Gruyter.

Khatri, V., & Brown, C. V. (2010). Designing data governance. Communications of the ACM, 53(1).

Laurent, A., Laurent, D., & Madera, C. (Ed.). (2019). Data lakes. ISTE Ltd / John Wiley and Sons Inc.

Lemieux, V. L., Gormly, B., & Rowledge, L. (2014). Meeting Big Data challenges with visual analytics: The role of records management. Records Management Journal, 24(2).

Reina, L. (2023). Noves arquitectures de dades. Lligall; Revista catalana d'arxivística, 46.

Serra Serra, J. (2024). El gobierno "archivístico" del dato. Tábula, 27.

Torreblanca, S. (2023). La governança de dades com a interacció: Un concepte analític per a les administracions públiques. Lligall; Revista catalana d'arxivística, 46.

Software

- Microsoft Power BI (desktop version)
- Microsoft Sharepoint
- Anaconda (Python)
- Orange Data Mining

Groups and Languages

Name	Group	Language	Semester	Turn
(TE) Theory	1	Catalan	first semester	afternoon