Information Systems and Data Science in the Humanities

Code: 45517
ECTS Credits: 6

2025/2026

Degree	Type	Year
Digital Humanities and Heritage	OP	1

Contact

Name: Juan Antonio Barceló Álvarez

Email: juanantonio.barcelo@uab.cat

Teachers

David Merino Recalde

Ermengol Gassiot Ballbe

Alessando Ravotto

Xavier Roda Gilabert

Teaching groups languages

You can view this information at the end of this document.

Prerequisites

No prior knowledge of computer science or programming is required, except for familiarity with computer equipment at an advanced user level. The required level of mathematics is that of compulsory secondary education.

Some familiarity with humanities and/or cultural topics is recommended.

English proficiency sufficient to read texts is required

Objectives and Contextualisation

The course introduces students to information management and Data Science, discussing the theory, techniques, and technology of relational databases and the most common structured query languages. It presents reference ontologies in the cultural domain (CIDOC-CRM) and their use in the indexing and cataloguing of cultural data. The application of these information systems in museums and archives is also discussed.
The course also introduces students to the use of Geographic Information Systems, statistical analysis, and data mining using machine learning techniques.

Learning Outcomes

1. CA13 (Competence) Analyse the limits and drawbacks of specific information management system designs.
2. CA13 (Competence) Analyse the limits and drawbacks of specific information management system designs.
3. CA14 (Competence) Explain the operation of digital information management systems that provide concrete solutions to problems arising from public use and open access.
4. CA15 (Competence) Research procedures for the management and processing of cultural and humanistic information with a gender perspective.
5. KA17 (Knowledge) Identify the most appropriate digital technologies for the indexing and cataloguing of humanistic and cultural information.
6. KA18 (Knowledge) Identify the operating principles and the most efficient statistical and/or machine learning (artificial intelligence) techniques for the processing of cultural and humanistic data.
7. SA19 (Skill) Apply digital text editing tools for semantic markup.
8. SA20 (Skill) Apply culture-specific ontologies, approved by UNESCO and other international organisations in the design of cultural information databases and management systems.
9. SA21 (Skill) Design systems for the computerised management of documents, and their indexing, cataloguing and consultation.
10. SA22 (Skill) Use statistical techniques, machine learning and data mining for data processing in the cultural and humanistic field.

Content

Introduction to databases and the relational model
Cultural data exchange. Metadata: Dublin Core, EUROPEANA (EDM), CIDOC-CRM, and others
Practical work with archaeological databases
Practical work with textual databases
Practical work with image databases
Geographic Information Systems.
Data analysis and statistics.

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Attendance at lectures led by the professor	18	0.72	CA14, CA15, KA17, KA18, CA14
Attendance at practical sessions with computer software led by the professor	18	0.72	CA14, CA15, KA17, KA18, SA19, SA20, SA21, SA22, CA14
Type: Supervised			
Classroom and Computer Lab practical work	34	1.36	CA13, CA14, CA15, KA17, KA18, SA19, SA20, SA21, SA22, CA13
Type: Autonomous			
Personal study. Bibliographical consultation. Additional	60	2.4	CA13, CA14, CA15, KA17, KA18, SA19,

practical work	SA20, SA21, SA22, CA13

Attendance at theoretical classes led by the professor.

Attendance at seminar sessions and practicals with computers and specific software led by the professor.

Classes are held in a specialized computer lab.

Comprehensive reading of texts.

The student is expected to make an independent effort to consult specialized bibliography. Part of the documentation is in English.

Independent practical computer exercises.

In-class debates, moderated by the teaching staff, on the most significant topics.

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

Continous Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Evaluation of practical work asked by the professor	40%	10	0.4	CA13, CA14, CA15, KA17, KA18, SA19, SA20, SA21, SA22
Presentation of a critical essay using Generative Artificial Intelligence	30%	5	0.2	CA13, CA15, SA21
Presentation of text commentaris upon references suggested by the professor	30%	5	0.2	CA13, CA14, CA15, KA17, KA18, SA21, SA22

The evaluation methodology for this master's course is based on the active and reflective participation of students. Their analytical skills will be assessed throughthe evaluation of practical exercises with computer software asked by the professor. Furthermore, comments on articles and bibliographic references will be also askes by the teaching staff. At the end of the course, students will prepare critical summaries of different technologies, expressing and arguingcriteria for good practice.

Another key component of the assessment will be a critical assignment involving the use of Generative Artificial Intelligence tools, applied to one of the topics discussed during the course. This assignment must include a reflection on the limitations and potential of these technologies within the field of Digital Humanities. The specific details regarding format, criteria, and deadlines will be explained and discussed in class by the professor.

Single assessment is not allowed.

At the time each assessment activity is carried out, the instructor will inform students (via Moodle) of the procedure and the date for reviewing grades.

Recovery procedure: only the final assignment (critical sumary) is eligible for reassessment. This decision will be made on a case-by-case basis following a personal interview between the student and the professor.

The submission date for reassessment will also be determined on a case-by-case basis and by mutual agreement between the professor and the student.

The student will receive a grade of "Not assessable" if they fail to submit any of the asked exercises.

If a student commits any irregularity that may significantly alter the grade of an assessment activity, that activity will be graded with a 0, regardless of any disciplinary proceedings that may be initiated. If multiple irregularities occur in the assessment activities of the same course, the final grade will be 0.

This course recommends the use of Artificial Intelligence (AI) technologies as an integral part of the development of assignments, provided that the final result reflects a significant contribution from the student in terms of analysis and personal reflection. The student must: (i) identify which parts were generated using AI; (ii) specify the tools used; and (iii) include a critical reflection on how these tools influenced the process and the final outcome of the activity.

A lack of transparency in the use of AI in this graded activity will be considered academic dishonesty and will result in a grade of 0 with no possibility of recovery, or more serious sanctions in severe cases.

Bibliography

Detailed references will be presented thorugh UAB Virtual Campus-MOOC.

Main general references:

Bruseker, G., Carboni, N., & Guillem, A. (2017). Cultural heritage data management: the role of formal ontology and CIDOC CRM. Heritage and archaeology in the digital age: acquisition, curation, and dissemination of spatial cultural heritage data, 93-131.

Burrough, P. A., McDonnell, R. A., & Lloyd, C. D. (2015). Principles of geographical information systems. Oxford university press.

Carlson, D. L. (2017). Quantitative methods in archaeology using R. Cambridge University Press.

Drennan, R. D. (2010). Statistics for archaeologists. New York: Springer.

Dritsou, V. (2024). Introduction to Cultural Heritage Data Modelling-with a focus on Europeana Data Model. DARIAH-Campus.

Foster, E., & Godbole, S. (2022). Database systems: a pragmatic approach. Auerbach Publications.

Hyvönen, E. (2012). Publishing and using cultural heritage linked data on the semantic web (Vol. 3). Morgan & Claypool Publishers.

Isaac, A., Fernie, K., Bachi, V., Tsoupra, E., Medici, M., Alkemade, H., ... & Heslinga, L. (2024). Making the Europeana Data Model a Better Fit for Documentation of 3D Objects. In 3D Research Challenges in Cultural Heritage V: Paradata, Metadata and Data in Digitisation (pp. 63-74). Cham: Springer Nature Switzerland.

Kennedy, M. (2009). Introducing geographic information systems with ARCGIS: a workbook approach to learning GIS. John Wiley & Sons.Oldman, D., & Labs, C. R. M. (2014). The CIDOC Conceptual Reference Model (CIDOC-CRM): primer. CIDOC-CRM official web site.

O'Neill, B., & Stapleton, L. (2022). Digital cultural heritage standards: from silo to semantic web. AI & society, 37(3), 891-903.

Otto, B., Ten Hompel, M., & Wrobel, S. (2022). Designing data spaces: The ecosystem approach to competitive advantage (p. 580). Springer Nature.Ranjgar, B., Sadeghi-Niaraki, A., Shakeri, M., Rahimi, F., & Choi, S. M. (2024). Cultural heritage information retrieval: past, present and future trends. IEEE Access.

Rockoff, L. (2021). The language of SQL. Addison-Wesley Professional.

Ruthven, I., & Chowdhury, G. G. (Eds.). (2015). Cultural heritage information: Access and management (Vol. 1). Facet publishing.

Silva, A. L., & Terra, A. L. (2024). Cultural heritage on the semantic web: The europeana data model. IFLA journal, 50(1), 93-107.

Yu, J. X., Chang, L., & Qin, L. (2022). Keyword search in databases. Springer Nature.

Software

MySQL. https://www.mysql.com/

MongoDB. https://www.mongodb.com/

QGIS. https://qgis.org/

PAST 5. https://www.nhm.uio.no/english/research/resources/past/

Groups and Languages

Please note that this information is provisional until 30 November 2025. You can check it through this link. To consult the language you will need to enter the CODE of the subject.

Name	Group	Language	Semester	Turn
(SEMm) Seminars (master)	1	Spanish	second semester	afternoon