Digital Technologies for Words and Sound

Code: 45513
ECTS Credits: 6

2025/2026

Degree	Type	Year
Digital Humanities and Heritage	OB	1

Contact

Name: Ramon Valdes Gazquez

Email: ramon.valdes.gazquez@uab.cat

Teachers

Maria Jesus Machuca Ayuso

Carlos Sanchez Lancis

Cecilio Garriga Escribano

Giuseppe Simone Pedote

Emanuele Leboffe

Lucia Cotarelo Esteban

Teaching groups languages

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

Prerequisites

To attend these studies, the general prerequisites of the MA degree on Humanities and Digital Heritage are necessary. In general, the student should have already some studies at BA-level on Humanities and / or Social Sciences disciplines. The course can also be useful to computer science graduates who want to specialize in the use of digital technologies in the field of Humanities and cultural studies, although they do not have previous experience on Humanities nor Cultural studies. Familiarity, at use level, with computers and standard office software is required.
The basic and reference bibliography is in English, as well as the software to be used. Knowledge of English at the level of specialized reading is therefore recommended.

Objectives and Contextualisation

This module aims to introduce students to the treatment and analysis of oral, written and sound productions with digital technologies. In the case of written texts and textual corpus, it is proposed to reflect on the implications of the transition from paper edition to digital edition and then focus on digital edition. In the case of oral and sound productions, an introduction will be made on the processing, labeling and categorization of sound files. The use of geographic information systems (GIS) for the coding of linguistic

information applied to the study of variation (geolinguistics) and the use of social networks and crowdsourcing as part of data mining will be explored. It will also reflect on the new conception of the literary text and its interpretation in the digital age, with special attention to the polysemic concept, as well as the new possibilities of approach to the artistic fact, the reception of digital artistic work in the field of the network and the new ways of approaching and analyzing the info-assisted text.

Learning Outcomes

1. CA11 (Competence) Use different multimedia and interactive technologies in a digital file with sound and text simply and efficiently.
2. CA12 (Competence) Interpret the digitisation and editing of text and sound for a better scientific understanding, teaching and dissemination of the humanities to different types of audiences.
3. KA13 (Knowledge) Describe the innovative potential of technology and the digital medium in the field of editing and access to texts and sound files.
4. KA14 (Knowledge) Apply the theoretical foundations of advanced digitalisation procedures to the design and implementation of a digital project.
5. KA15 (Knowledge) Incorporate advanced-level computerised management strategies for text or sound files into a digital project.
6. KA16 (Knowledge) Report on the use of multimedia technologies and AI-based approaches to increase the accessibility and communicability of the digitisation of humanistic and cultural data and computer vision programs.
7. SA15 (Skill) Analyse the practical problems arising from the application of digitisation and computer vision in the field of humanities and cultural studies.
8. SA16 (Skill) Solve practical problems by digitally editing text and/or sound files and using markup languages.
9. SA17 (Skill) Classify digital texts by their subject matter and content of computer tools to help in their cataloguing and indexing.
10. SA18 (Skill) Analyse textual, linguistic, literary, historical-cultural, and sound data through the use of statistical techniques, digital tools, machine learning, and data mining.

Content

Digitization and recognition of written documents. Scanning and OCR.

From analog speech signals to digital signals.

Segmentation, Labeling and Storage.

WAV and MP3 audio formats.

Linguistic and literary data extraction, statistical analysis, and inference.

Publication on audio document networks and text grids.

Terminology management and research, and relational databases.

From manuscripts and printed texts to XML TEI: digital edition.

Segmentation, markup, and analysis of literary texts and linguistic corpora with XML TEI.

TEI modules, literary genres, and Relax rng schemas.

New forms of research and dissemination in Literature: bases of data, magatzematge, visualization and analysis of data.

Activities and Methodology

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Problem-based learning. Case-based learning. Classroom practical work. Seminars. Workshops. Debates.	26	1.04	CA11, CA12, KA13, KA14, KA15, KA16, CA11
Theoretical classes with an explanation of computer techniques and their theoretical and methodological foundations	40	1.6	CA11, CA12, KA13, KA14, KA15, KA16, SA15, SA16, SA17, SA18, CA11
Type: Supervised			
Presentation of computer equipment.	13	0.52	KA16, SA18, KA16
Search for documentation, elaboration of databases, digital editions, exercises of application of the presented analysis and study techniques, reading of texts, writing of works.	28	1.12	CA11, CA12, KA13, KA14, KA15, KA16, SA15, SA16, SA17, SA18, CA11
Type: Autonomous			
Practical work with hardware and software.	23	0.92	CA11, KA14, KA16, SA15, SA16, CA11

The methodology is divided between directed activities, supervised activities, autonomous activities and assessment activities.

In autonomous activities (22,4%), study hours and student preparation must be taken into account in order to face the assessment activity. These activities will be composed of searching for documentation, elaboration of databases, exercises to apply the exposed study techniques and reading references as reinforcement material.

The directed activities (48,8%) have to respond in a predetermined time schedule, which requires the face-to-face address of a teacher and which is specified in hours in the previous section. In addition, it must be taken into account that these activities are distributed in theoretical classes (28,8%) and approach to case studies and problems that may arise around a specific topic (20%).

Regarding supervised activities (28,8%), the teacher programs them so that the student works autonomously, but with the teacher's supervision. In case the student cannot develop these activities autonomously, the teacher will suggest the materials that he can use to carry out the proposed activities.

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.

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
Presentation of a critical essay using Generative Artificial intelligence	40%	10	0.4	CA11, CA12, KA13, KA14, KA15, KA16, SA15, SA16, SA17, SA18
Presentation of a written essay debating some relevant subject	30%	5	0.2	CA12, KA13, KA16, SA15, SA17
Presentation of written essais commenting some text	30	5	0.2	CA12, KA13, KA14, KA15, KA16, SA15, SA16, SA17, SA18

The evaluation methodology for this master's course is based on the active and reflective participation of students. Their analytical skills will be assessed through written comments on articles and bibliographic references proposed by the teaching staff. In addition, students will prepare critical summaries of class debates, where they must express and argue their own position in relation to opposing opinions, demonstrating their ability for dialogue and critical thinking. 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 allowed, upon request, following the procedure approved by the Dean of the Faculty of Arts.

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 (second case study) 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-casebasis 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 obligatory essays.

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

Allés, S., Materiales en Zenodo para el aprendizaje de TEI. 2020. Open Access, recurso en línea: https://tthub.io/aprende/

Amores, M., "El buscador GICES XIX, herramienta digital sobre el cuento español del siglo XIX", en López Poza, S.; Pena Suerio, N., *Humanidades digitales. Desafíos, logros y perspectivas de futuro*, 2014.

https://dialnet.unirioja.es/servlet/libro?codigo=660935

Audio and Video Guidance: Resources, U.S. National Archives. https://www.archives.gov/preservation/formats/audio-video-resources

Bel, N. 2014. Documentation of ContaWords. http://hdl.handle.net/10230/22602

Buenafuentes, C. y Sánchez Lancis, C. (en prensa): "The Spanish Twenty-first Century Corpus (CORPES XXI): a Tool for the Study of Syntactical Variation in Spanish", en Cerrudo, A.; Gallego, Á. y Roca, F. (eds.),

Syntactic geolectal variation: traditional approaches, current challenges and new tools. Amsterdam: John Benjamins.

Cerrudo, A. et al. 2015. "ASinEs: Prolegómenos de un atlas de la variación sintáctica del español". *Linguamatica* 7.2.: 59-69. https://linguamatica.com/index.php/linguamatica/article/view/V7N2.5

Cobo. Á., *PHP y MySQL: Tecnología para el desarrollo de aplicaciones web*. Madrid, Ediciones Díaz de Santos, 2005.

Collatón, R., *Introducción al uso de R y R Commander para el análisis estadístico de datos en ciencias sociales*. Comunidad de programadores, 2014. Extraído de https://cran.r-project.org/doc/contrib/Chicana-Introduccion_al_uso_de_R.pdf

Correa Duarte, J.A., *Manual de analisis acustico del habla con Praat*. Publicaciones del Instituto Caro y Cuervo, Series Minor 49, Bogotá, 2014.

Dickinson, M.; Brew, Ch.; Meurers, D. Language and computers. London: Wiley-Blackwell, 2012.

Driscoll, M.J., y Pierazzo, E., eds., *Digital Scholarly Editing. Theories and Practices*, Open Book Publishers, Cambridge (UK), 2016. Recurso en línea, descarga gratuita. http://dx.doi.org/10.11647/OBP.0095

Franzini, G., *Catalogue of Digital Editions*, UCL Centre for Digital Humanities y Austrian Centre for Digital Humanities, 2012-... recurso en línea:https://dig-ed-cat.acdh.oeaw.ac.at/

Huidobro, J.M. "Sonido digital y formatos de compresión", *Acta 24*. 2002. Extraído de https://www.acta.es/recursos/revista-digital-manuales-formativos/358-024

International Association of Music Libraries, Archives and Documentation Centres: https://www.iaml.info/

Marrero, V. (Ed.), *Introducción a la fonética judicial. Variación inter e intralocutor en español*. *El proyecto VILE*, Tirant lo Blanch, Valencia, 2017.

Pierazzo, E., *Digital Scholarly Editing: Theories, Models and Methods*, Farnham (Surrey),Ashgate Publishing, 2015.

Puertas Pavón J., *Creación de un portal con PHP y MySQL*, RA-MA, S.A. Editorial, Madrid, 2015.

Sahle, P.,*A Catalogue of Digital Scholarly Editions*, Institut für Dokumentologie und Editorik, Universität zu Köln, Colonia, 2008-..., recurso en línea: http://www.digitale-edition.de/

Sound Directions: Best Practices for Audio Preservation, Indiana University Digital Library Program. http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/index.shtml

Software

Gephi (graphs): < https://gephi.org/ >.
Onodo (redes): < https://onodo.org/ >.
Oxygen (Editor XML): < https://www.oxygenxml.com/ >.
R (Lenguaje de programación): < https://www.r-project.org/ >.
RStudio (Entorno para R): < https://posit.co/download/rstudio-desktop/ >.
Stylo (Librería para R): < https://eadh.org/projects/stylo-r-package >.
Timemapper (cronologías, timelines y mapas): < https://timemapper.okfnlabs.org/ >.
Transkribus (Versión web): < https://www.transkribus.org/ >.
Visual Studio Code: < https://code.visualstudio.com/ >.

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	Catalan/Spanish	first semester	afternoon