ADVANCED METHODS FOR LITERATURE ANALYSIS: BIBLIOMETRICS, COMPUTATIONAL REVIEWS AND GENERATIVE AI TOOLS

Dr. Michele Cipriano, Postdoctoral Researcher
Faculty of Economics and Law, Department of Economic and Social Sciences UCSC-Piacenza

Dr. Marco Smacchia, RTD-A
Department of Economics, Engineering, Society and Business Organization University of Tuscia

Course Aims

The course aims to equip doctoral students with theoretical and practical competences for conducting rigorous and reproducible literature analyses through advanced quantitative and computational methods. Recognising that scientific knowledge progresses cumulatively, the course focuses on how to structure, map, and interpret existing literature in a transparent and data-driven way. Students will learn how to apply bibliometric analysis, topic modelling, and generative AI tools (LLMs) to explore the conceptual, intellectual, and social structures of a research field, with particular reference to agrifood system studies.

At the end of the course, participants will be able to:
- Design and conduct a systematic literature review using bibliometric data;
- Collect and preprocess bibliographic metadata from academic databases;
- Perform descriptive and science mapping analyses with Bibliometrix and Biblioshiny;
- Apply computational techniques (e.g., BERTopic) to identify latent themes and trends;
- Integrate Generative AI tools (e.g., ChatGPT, Gemini) to support interpretation, synthesis, and writing;
- Apply these methods to their own doctoral research topics in agrifood and sustainability systems.

Methodology

The course combines lectures, live demonstrations, and hands-on sessions. Each meeting includes a theoretical introduction, tool demonstration, and guided exercises using real bibliometric datasets and student-selected topics.

Students are required to bring a personal laptop (minimum 8 GB RAM) and ensure a stable Internet connection.
The following software and accounts are required:
- R and RStudio (for Bibliometrix and Biblioshiny);
- Google account (for shared materials and Google Colab);
- Access to a Generative AI tool (ChatGPT, Claude, Gemini, or similar; free version sufficient).

No programming background is required. The course is designed to be inclusive and accessible to students from various disciplinary backgrounds.

Course description

Session 1 (4h) – Bibliometric Analysis: Concepts and Applications (by Dr. Cipriano)

This session introduces the conceptual and methodological foundations of bibliometric analysis, grounded in Graph Theory and Social Network Analysis (SNA). Students will learn to collect and manage bibliographic data, and to conduct descriptive and science mapping analyses with Bibliometrix and Biblioshiny.

Key topics:
- Graph theory and network measures;
- Metadata collection and query design;
- Bibliometric indicators and visualization;
- Conceptual, intellectual, and social structures of the literature;
- Bibliographic coupling (co-citation, co-occurrence, co-word).

Lab: Practical exercises using Biblioshiny to build and interpret bibliometric networks.

Session 2 (4h) – Computational Literature Reviews and Topic Modelling (by Dr. Smacchia)

This session focuses on computational approaches to large-scale literature analysis and topic discovery. Students will explore how to combine quantitative mapping with qualitative interpretation to generate deeper insights and theoretical contributions.

Key topics:
- Computational literature review (CLR) and text mining fundamentals;
- Topic modelling with BERTopic and other algorithms;
- Integration of bibliometric and topic modelling outputs;
- Using LLMs (e.g., ChatGPT, Claude) to support thematic synthesis and data interpretation;
- Reflections on theory building from computational results.

Lab: Application of a topic modelling workflow to a sample corpus, followed by interactive discussion.

Session 3 (4h) – Final Assignment and Guided Project (by Dr. Cipriano & Dr. Smacchia)

Students will develop a small-scale research project based on one of the introduced methods. Working on a topic of personal research interest, they will:
- Define a research question;
- Collect bibliometric data;
- Apply one analytical method (bibliometric, CLR, or LLM-based);
- Produce a short report (2–3 pages) summarizing the methodology, findings, and reflections.

Reports must be submitted within 10 days after the last session. Feedback will be provided by instructors during a final discussion.

Recommended texts

Scientific papers, slides, lecture notes, and the textbook reference provided by the lecturers