COMP 7980: Capstone Project
Spring 2027

Digital Humanities
Capstone Project, Spring 2027

If you sign up for CS Capstone, COMP 7980, you'll have the opportunity to work with project director Dr Beverley Rilett in Auburn University's Special Collections and Archives. Individually or with a group, you will select one of the exciting projects below to complete during the semester.

COMP 7980 MSc Capstone: Building Digital Tools for Humanities Research

This course offers computer science students the opportunity to engage in interdisciplinary, project-based learning at the intersection of technology and the humanities. Working with Dr. Beverley Rilett in Auburn University's Special Collections and Archives, students will collaborate on creating innovative digital tools and applications used by researchers worldwide in the fields of Victorian Studies and Digital Humanities. Projects involve applying advanced technologies such as AI, text-to-speech, sentiment analysis, and data visualization to analyze and present historical and literary data. Through hands-on, team-based work, you will gain practical experience in software development, data processing, and digital resource design while contributing to meaningful, real-world research initiatives.

Key Details

- **Weekly meetings**: We'll meet once per week in person at the library, likely in a group with others working on similar projects. These meetings are designed to foster collaboration, friendship, and creativity as you develop real-world applications used by researchers globally. You will work on these applications remotely, with your team members and individually, for approximately 6-8 hours per week.

- **Scheduling**: Send me your course schedule so we can find a one-hour block of time to meet in person each week, at a time that works for everyone. You do not have to schedule other classes around this one. We'll figure it out.

Project Overview: Computational Realism: Leveraging LLMs for the Analysis and TEI Optimization of George Eliot's Novels

This project explores the intersection of Victorian-era literature and natural language processing (NLP) by applying large language models (LLMs) to the works of George Eliot. Eliot's novels, characterized by complex psychological realism, intricate social networks, and philosophical interventions, provide a rich dataset for computational analysis.

You will utilize LLMs (including Claude Code) to perform advanced semantic analysis, thematic extraction, and character network mapping across Eliot's major works (e.g., *Middlemarch*, *Daniel Deronda*, and *The Mill on the Floss*).

You will use LLMs to automate and refine the tagging of metadata, structural elements, and named entities within TEI-XML frameworks, ensuring a higher degree of granularity and accuracy than traditional manual encoding allows.

COMP 7980: Capstone Project
Spring 2027

You will create data visualization tools to help researchers of George Eliot's life and works understand thematic, chronological, geographical, emotional, and intellectual patterns across her writings. The project's editorial team will provide the cleaned datasets; developers completing COMP 7980 will apply computational analysis to find new ways of interpreting the datasets.

Expected Outcomes:

1. **Enhanced Digital Corpus:** A set of optimized TEI-XML files for George Eliot's novels, featuring improved semantic tagging, character identification, and structural consistency.

2. **Methodological Framework:** A documented workflow for using LLMs to clean, update, and enrich legacy TEI files in the digital humanities.

3. **Literary Analysis Report:** A comparative study demonstrating how LLM-driven insights (such as sentiment shifts or social-strata vocabulary) align with or challenge traditional literary critiques of Eliot's realism.

4. **Data Visualizations:** Character interaction maps and thematic heatmaps derived from the LLM-processed data, illustrating narrative patterns across the George Eliot corpus.