



VIA
PANORAMICA

Revista de Estudios Anglo-Americanos
A Journal of Anglo-American Studies

TOWARDS THE MAELSTROM THE DIGITAL HUMANITIES IN ACTION

1ST CETAPS MEETING ON
DIGITAL HUMANITIES
DOSSIER

THEMATIC SECTION

EXTRA

Looking From a Distance: An approach to text analysis with R

Luciano Moreira, Felipe Ferreira, Joana Pinela, Joel Faustino,
Raquel Souza, Lile Kobaliani, Thaís Barbedo Vigas
& Lizi Kolbaia
University of Porto

ABSTRACT: In this workshop, we explored a research example on the field of utopian studies and, within it, feminist criticism to illustrate text analysis with R language using metadata from the Lyman Tower Sargent Bibliography. Participants were introduced to the research topic, the R Studio environment, and guided through the iterative process of text analysis (word frequency and network analysis) with the Quanteda library. We aimed to show that visualizations are, more than results, part of a process that triggers new, more complex research questions.

KEYWORDS: R Language, Utopian Studies, Feminist Criticism, Text Analysis

RESUMO: Neste workshop, explorámos um exemplo de pesquisa no campo dos estudos utópicos e, dentro dele, da crítica feminista para ilustrar a análise de texto com linguagem R, usando metadados da base de dados “Lyman Tower Sargent Bibliography”. Os participantes foram introduzidos ao tema de pesquisa, ao R Studio, e orientados pelo processo iterativo de análise de texto (frequência de palavras e análise de rede) com o recurso à Quanteda Library. O nosso objetivo foi demonstrar como as visualizações são, mais do que resultados, parte de um processo que desencadeia novas e mais complexas questões de investigação.

PALAVRAS-CHAVE: Linguagem R, Estudos Utópicos, Crítica Feminista, Análise de Texto

INTRODUCTION

The 1st CETAPS Meeting on Digital Humanities brought together scholars to explore how computational methods are shaping academic research. Among the highlights of the event was the Digital Lab and Junior Researchers in Anglo-American Studies (JRAAS) team's workshop, which explored the potential of digital tools to analyze the structure and development of academic disciplines. Using feminist utopian literature as a case study, the workshop demonstrated how digital humanities may uncover thematic trends and structural imbalances within academic fields. This effort drew on the previous study "Looking From A Distance: Digital Humanities Insights into the Field of Feminist Utopian Criticism" (Moreira *et al.* 2024) which was presented at the "2nd International Conference on Data & Digital Humanities", at the University of Minho in December 2024.

The primary dataset for the workshop was the Lyman Tower Sargent Bibliography,¹ a collection of over 20,000 references curated by CETAPS' ARUS project. This bibliography, widely recognized as a critical resource in utopian studies, spans reference works from across the globe and several decades. However, it also reveals striking imbalances. Roughly 70.6% of its entries are in English, underscoring a lack of representation for non-English-speaking scholars. Moreover, feminist themes are underrepresented despite their significant contributions to the field. These disparities set the stage for the workshop, which used digital tools to investigate how such imbalances shape the construction of academic knowledge.

Feminist utopian literature, which was at the heart of the workshop, is a genre that critiques patriarchal systems while imagining alternative futures. This body of work provides a space for reflecting on societal inequalities and envisioning transformative possibilities. The metadata offered by the ARUS repository represented the genre's ability to explore intersections of gender, race, and class. Yet despite its richness, feminist utopian literature is yet to receive the academic attention it deserves. The genre's underrepresentation reflects broader systemic biases within scholarly research, a topic the workshop addressed by analyzing how these patterns manifest in bibliographic datasets.

A key method introduced during the workshop was distant reading, a concept developed by literary scholar Franco Moretti (2000). Unlike close reading, which involves analyzing individual data in detail, distant reading focuses on identifying larger patterns across extensive datasets. This approach fits within the broader goals of digital humanities,

which emphasize not just digitization but also critical reflection on what is digitized and how it is studied (Fitzpatrick 2012; Gold & Klein 2016). Participants applied distant reading to the Lyman Tower Sargent Bibliography, using computational methods, to examine thematic trends and identify systemic biases.

IMPLEMENTATION

The workshop offered participants hands-on experience with R, a programming language frequently used in digital humanities for data analysis and visualization. Attendees used tools such as `quanteda`, `dplyr`,² and `ggplot2` to work with a subset of the bibliography, focusing on feminist utopian literature. This subset included 453 works and 3,062 tokens. Through activities like tokenization, data visualization, and statistical analysis, attendees identified thematic patterns and relationships within the data. Outputs such as word clouds, frequency tables, and network diagrams provided a visual representation of the genre's contributions to utopian studies.

The analysis revealed several noteworthy trends. Feminist themes were connected to broader topics like science fiction, social critique, and future-oriented narratives. These findings affirmed the centrality of feminist utopian literature within utopian studies while also highlighting the systemic barriers that limit its recognition in the academic field. For example, despite its thematic richness, feminist literature occupies a relatively modest share of the bibliography, raising questions about how academic disciplines have historically marginalized certain voices. These discussions resonated with Moretti's (2016) call for digital humanities to challenge traditional research methods and expand the scope of scholarly inquiry.

The workshop's interactive format allowed participants to engage deeply with both digital tools and the theoretical questions they raised. Activities included importing and cleaning data, running statistical analyses, and generating visualizations to illustrate the dataset's structure. For instance, participants created network diagrams that mapped thematic relationships and word clouds that highlighted frequently occurring terms. These exercises not only introduced participants to practical techniques but also prompted them to reflect on how digital tools can reveal insights that might otherwise go unnoticed.

Alongside technical training, the workshop encouraged critical discussions about the implications of using digital methods in research. Participants explored challenges such as algorithmic bias, the limitations of existing datasets, and the complexity of combining quantitative and qualitative approaches. For example, feminist themes, as a minority perspective in the dataset, prompted debates about how different approaches to metadata could enhance their visibility. This dialogue underscored the potential of digital humanities to reshape the ways in which academic fields are represented and studied.

CONCLUSIONS

One of the workshop's key insights was the ability of digital tools to uncover biases embedded in academic repositories. The relatively small representation of feminist topics in the Lyman Tower Sargent Bibliography reflects larger systemic inequalities within the academic community. By addressing these imbalances, the workshop demonstrated how digital methodologies can challenge conventional narratives and make room for a more inclusive and representative academic landscape.

While the workshop focused specifically on feminist utopian literature, its findings have broader implications for understanding how academic disciplines are constructed. By analyzing large datasets, computational tools like R allow researchers to identify patterns, uncover imbalances, and reflect critically on the factors shaping academic fields. These insights can inform efforts to create more equitable and inclusive scholarly practices.

As the session drew to a close, participants discussed how they could apply what they had learned to their own research. Topics ranged from analyzing bibliographic data to exploring thematic trends and questioning structural inequities in their respective fields. The workshop combined theoretical exploration with practical application, leaving attendees better equipped to engage with digital methodologies in meaningful ways.

The workshop at the "1st CETAPS Meeting on Digital Humanities" was an important step in advancing the dialogue around computational methods in scholarship. By focusing on feminist utopian literature and its broader academic context, the workshop offered a compelling mix of technical training and critical reflection. As digital humanities continue to grow, initiatives like this will play a vital role in shaping how we study and understand academic disciplines, ensuring that new methods not only expand knowledge but also challenge rooted inequalities.

END NOTES

¹ <https://cetapsrepository.letras.up.pt/id/cetaps/94026>

² <https://dplyr.tidyverse.org/>

REFERENCES

- Fitzpatrick, K (2012) *The humanities, done digitally*. K Fitzpatrick. *Debates in the digital humanities*, 12-15.
- Gold, M. K., & Klein, L. F. (Eds.). (2016). *Debates in the Digital Humanities 2016*. University of Minnesota Press. <https://doi.org/10.5749/j.ctt1cn6thb>
- Moretti, F. (2000). “Conjectures on world literature”. *New Left Review*, (1), 54–68.
- Moretti, F. (2013). “Distant reading”. London: Verso.
- Moreira, L., Pinela, J., Souza, R., Faustino, J., Ferreira, F., Kobaliani, L., & Benison, L. (2024). “Looking From A Distance: Digital Humanities Insights into the Field of Feminist Utopian Criticism”. *2nd International Conference on Data & Digital Humanities: Generative Artificial Intelligence for Text and Multimodal Data*. Braga. (Paper.)

HOW TO CITE

Moreira, Luciano, Felipe Ferreira, Joana Pinela, Joel Ramos Faustino, Raquel Souza, Lile Kobaliani, Thaís Barbedo Vidas & Lizi Kolbaia (2024). “Workshop 2: Looking from a Distance: An Approach to Text Analysis with R”. *VIA PANORAMICA: Revista de Estudos Anglo-Americanos*, vol. 13, n.º 2, 2024, pp. 99-103. Web: <http://ojs.letras.up.pt/>. DOI: 10.21747/2182-9934/via13_2w2