

Looking From A Distance: Digital Humanities Insights into the Field of Feminist Utopian Criticism

Luciano Moreira^[0000-0003-3162-1021], Joana Pinela^[0009-0001-2659-5191], Joel Faustino^[0009-0004-0257-0805], Raquel Souza^[0009-0007-7869-8065], Felipe Ferreira^[0009-0005-0646-8186], Lile Kobaliani^[0009-0001-0913-9346], Liam Benison^[0000-0002-4989-4370]

CETAPS - Centre for English, Translation, and Anglo-Portuguese Studies, Faculty of Arts and Humanities of the University of Porto, 4150-564 Porto, Portugal
lucianomoreira@letras.up.pt

Abstract. In this paper, we aimed to explore how digital humanities can contribute to a better understanding of the construction of scientific disciplines. We specifically focused on the field of utopian studies and within it, feminist criticism, using metadata from the field's largest bibliographic database. The Lyman Tower Sargent Bibliography includes over 19,000 critical references, around 70% of which are written in English. We conducted the statistics and text analysis in R. Results showed that works mostly address literature. Women-related topics, despite representing a relatively small part of the corpus, rank as the third most important topic and are integrated into the most complex thematic constellation, which connects, for example, the themes of science fiction and the future. This work is significant because it shows that digital humanities can help understand heuristics, imbalances, and trends in the construction of repositories and the fields they aim to represent.

Keywords: Digital humanities, Utopian literature, Feminist criticism.

1. Introduction

The digital humanities (DH) have transformed research practices, with digital technology mediating these processes, while the field continually evolves and redefines itself (Berry, 2012; Callaway et al., 2020). This ongoing transformation makes defining digital humanities challenging due to its interdisciplinary and expansive nature (Gold & Klein, 2016), and researchers are particularly concerned with avoiding the perception of the field as mere digitization (Fitzpatrick, 2012). Thus, it is important to emphasize that DH transcends mere digitization, as its focus is not solely on the process of digitization itself, but rather on the critical questions of what should be digitized, the purposes behind digitization, and the methodologies

employed. This approach facilitates a deeper understanding of corpora and the evolution of scientific disciplines. In line with this broader focus, digital humanities have introduced innovative approaches such as distant reading, which enables the examination of large-scale literary patterns, reshaping how texts can be analyzed and understood.

Thus, the objective of our research is to explore the contribution of the Digital Humanities to the advancement of our understanding of the construction of scientific disciplines by focusing on feminist utopianism. To answer it, we will analyse the Lyman Tower Sargent Bibliography, an extensive collection of critical literature on utopia. We will actively retrieve and revise the relevant information from the dataset - thus turning data into capta (Drucker, 2011) - eventually focusing our analysis on critical feminist literature on utopia. An analysis such as this - employing the method of distant reading of metadata- can offer a uniquely broad view of the field of feminist utopianism that can inform future writings on the topic, by detecting trends and exceptions.

2. Literature Review

2.1. Distant reading

The term “distant reading” was coined by Franco Moretti in 2000. This approach allows analysis of units much smaller or larger than the text, such as devices, themes, and genres (Moretti, 2000, pp. 48-49). Despite his role in establishing DH, Moretti (2016) acknowledges the difficulty in defining the field, admitting that “digital humanities means nothing”. DH should not be seen as the savior of the humanities; instead, the humanities must assert their significance (Moretti, 2016). Recently, large-scale databases like IMDb and TV Tropes have become prevalent, opening new analytical methodologies and raising challenges that DH can explore. This growing number of databases implies a better understanding of how to preserve and share data in the digital era. This critical thinking of data leads to the concept of capta proposed by Drucker (2011). According to her, data “is assumed to be given” and capta, is information captured to fit an interpretative process, suggesting that capta is not a neutral or self-evident reflection of reality. As the DH continue to evolve, the potential for exploring broader theoretical concepts, such as utopian visions, becomes increasingly apparent, illustrating how these innovative methodologies can uncover new dimensions in literary studies.

2.2. Digital Humanities

DH has a utopian core as expressed in the manifesto *The Digital Humanities Manifesto 2.0* ([Reference Deleted for Blind Review (RDFBR)], 2024). Brian Greenspan (2016) argues that when utopias are linked to DH, they typically represent a vision of a promising future with expanded opportunities and openness. This utopian core is also shared by the ARUS database analyzed in this paper. A distant reading of collections like the Lyman Tower Sargent Bibliography can provide insights into utopian studies that traditional humanities could not achieve without digital tools. Analyzing this extensive literature allows us to reevaluate and deepen our understanding of utopias and utopian thought.

2.3. Utopian studies

The Lyman Tower Sargent Bibliography, managed by Advanced Research in Utopian Studies (ARUS) at the Centre for English, Translation and Anglo-Portuguese Studies (CETAPS), is the leading bibliographical repository of critical literature on utopianism, containing over 19,000 works. The project relies on contributions from scholars interested in adding digital objects to the database. Its (rather utopian) aim is to compile a comprehensive list of critical sources in utopian studies, in any language, as abstracts or full texts, allowing researchers to easily browse by organizing sources by subjects.

As its name indicates, the bibliography was organized by Lyman Tower Sargent, a North American academic and professor emeritus at the University of St. Louis, Missouri, and the foremost expert on utopian studies in the world. Sargent (1994) defined utopianism as “social dreaming,” where imagination envisions radically different societies rooted in social realities. He categorized it as having three “faces”: literary utopias (further divided into body and city utopias), communitarianism, and utopian social theory, each representing different expressions of human social reality. However, scholars such as Gregory Claeys (2012) wrote against reductivity in defining utopia, arguing that no single definition would ever be satisfactory; accepting the broad validity of Sargent’s categorization of utopianism, Claeys argued that it nevertheless needed reordering, proposing a new “social realist” definition. Ruth Levitas (2013) argued that utopian studies and sociology developed alongside each other and that they need to come together for a better understanding of both, favoring a definition of utopia as being “a repressed dimension of the sociological”. To further explore the topic, this paper focuses on feminist utopianism and its representation in the Lyman Tower Sargent Bibliography.

Defining feminist utopianism is as challenging as defining utopianism itself. Sarah Lefanu describes feminist utopias as “imaginary places, a nowhere land, a realm like the unconscious” (1985), which allow feminists to escape “the restrictions of patriarchal scholarship” (Sargisson, 1995). Feminist utopias seem to have evolved

diachronically, reflecting the changing priorities, challenges, and rifts of the feminist movement over time. Feminist utopian literature united by the ARUS database demonstrates this process of historicization.

2.4. Feminist utopianism

Historically, feminist utopias were overlooked until the 1970s when feminist criticism began emphasizing women's writing. Early examples include Christine de Pizan's *The Book of the City of Ladies* (1405), Margaret Cavendish's *The Description of a New World Called the Blazing World* (1666), and Mary Griffith's *Three Hundred Years Hence* (1836) (Fancourt, 2004). Emerging as early as the fifteenth century, these texts critiqued societal norms and power structures, laying the foundation for feminist discourse by providing a literary space to explore women's liberation and social justice.

Feminist utopias of the early twentieth century such as Carol Elizabeth Corbin's *The Position of Women in the Socialist Utopia* (1901) and Charlotte Perkins Gilman's *Herland* (1915) focused on achieving gender equality through reform and education.

In the second wave, from the 1960s to the 1980s, feminist utopias expanded to address sexual liberation, communal living, and economic independence, reflecting the diverse ideas of the women's liberation movement (Fancourt, 2004). Texts such as Monique Wittig's *Les Guérillères* (1969), Joanna Russ's *The Female Man* (1975), Marge Piercy's *Woman on the Edge of Time* (1976), and Sally Miller Gearhart's *The Wanderground* (1978) envisioned worlds where traditional gender roles were dismantled and alternative social structures were explored (Marketos, 1993).

Third-wave feminist utopias from the 1990s to the early 2000s reflect a more nuanced understanding of identity, emphasizing the intersection of gender with race, class, sexuality, and other axes of identity. These texts envisioned diverse and inclusive societies, as seen in Marge Piercy's *He, She and It* (1991), Octavia Butler's *Parable of the Sower* (1993), and Larissa Lai's *Salt Fish Girl* (2002). Contemporary feminist utopias, from the 2010s to the present, incorporate the impact of technology and digital spaces on gender relations, alongside concerns for environmental sustainability. Notable examples include Louise Erdrich's *Future Home of The Living God* (2017), Lauren Beukes's *Afterland* (2020), and Christina Sweeney-Baird's *The End of Men* (2021).

3 Methods

To systematically explore these diverse feminist utopian narratives, this study employs a detailed methodology, to access and analyze a vast collection of feminist utopian literature through systematic data retrieval and statistical analysis.

3.1 Data

The CETAPS Repository is a database containing more than 30,000 items. Built with DSpace, an open-source repository software package, this database is an adequate solution for preserving and accessing digital data. The backend in SQL allows relational storage of data and the use of Apache Solr permits powerful and rapid searches throughout the database.

The Lyman Tower Sargent Bibliography of the University of Porto is one of the most important collections in the CETAPS Repository. With 19,604 items - dated from 1625 to 2023 – and still growing (the last update being in June 2024), this collection is composed of works dating from 1625 to 2023. The items are classified according to 28 different metadata fields, including contributor, language, publication title, and date. As for the language, 70.6% of them were written in English; other languages had an expression below 30%. As for document type, journal articles constituted 44.2% of all entries, and book chapters, 39.5%.

3.2 Procedures

We retrieved the Lyman Tower Sargent database from the CETAPS Repository and imported it into an R environment, utilizing different libraries such as `quanteda`, `dplyr`, and `ggplot2`, to produce different analyses and visualizations.

We filtered the general data and created a subset composed of all the English-language works in the repository. The titles were then broken up into tokens (Table 1).

Table 1. Lyman Tower Sargent Bibliography: Works and tokens

Corpus	No. of works	No. of tokens	Mean token no.	Standard deviation
General	19,604	108,071	6	3
English-language subset	13,839	73,186	5	3
Feminist theme subset	453	3,062	7	3

The tokens were subsequently organized by themes in the relevant cases. For example, ‘utopia’, ‘utopias’, ‘utopianism’, and ‘utopian’ were all placed within a category called ‘utopia’. Another example is the category ‘women’, which includes tokens such as ‘woman’, ‘women’, and ‘female’. We also grouped tokens that composed specific names of relevant authors and works, such as ‘margaret’ and ‘atwood’.

Out of this data collection, we ran several analyses, producing several outputs: token frequency tables, token chronological distribution graphs, quantile graphs, and token relational networks.

4. Results

In this section, we can describe the main results. After analyzing the complete database to obtain some general information as described in the methods, we created an English subset corpus with 13,829 works. We divided the distributions of papers per year into 10 quantiles. Through this analysis, we observed that there was little focus on utopian critical production between 1652 and 1962. Indeed, we could verify that the peak of production happened in the early twenty-first century between the years 2000 and 2014.

The 50 most frequent words in the titles are shown in **Fig. 1**. We observed that the most frequent words/themes are: ‘utopia’ and ‘science fiction’. Within the ten most frequent words we find words related to feminism such as “women”, which occupies the third position. However, we do not find any female authors; on the contrary, we find male authors such as Thomas More and William Morris.

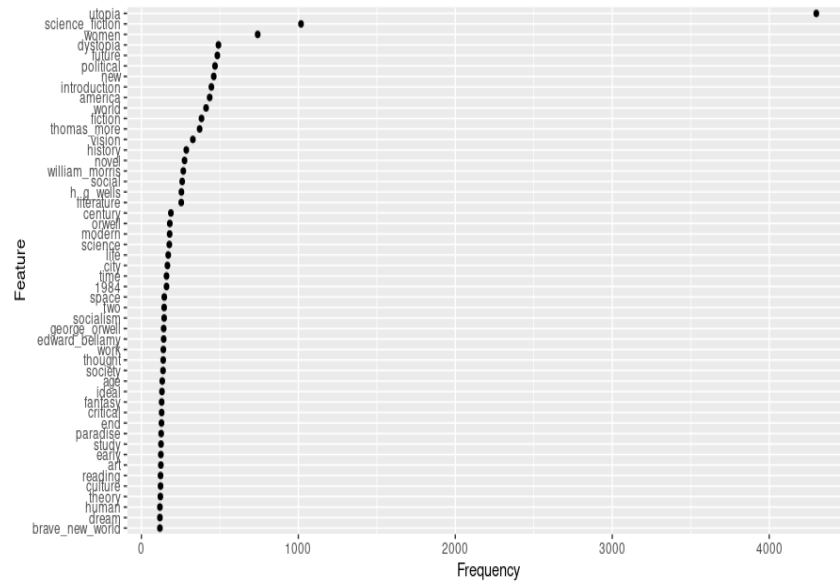


Fig. 1. Frequency of the words/themes English subset

We can observe in **Fig. 2** the connections between utopia and science fiction. ‘America’ and ‘Thomas More’ also have a strong connection with the word ‘utopia’. However, we can observe that the women-related words (the theme ‘women’) are placed in the most complex constellation of the map, connecting with ‘utopia’, ‘science fiction’, ‘new’, ‘fiction’, and ‘future’.

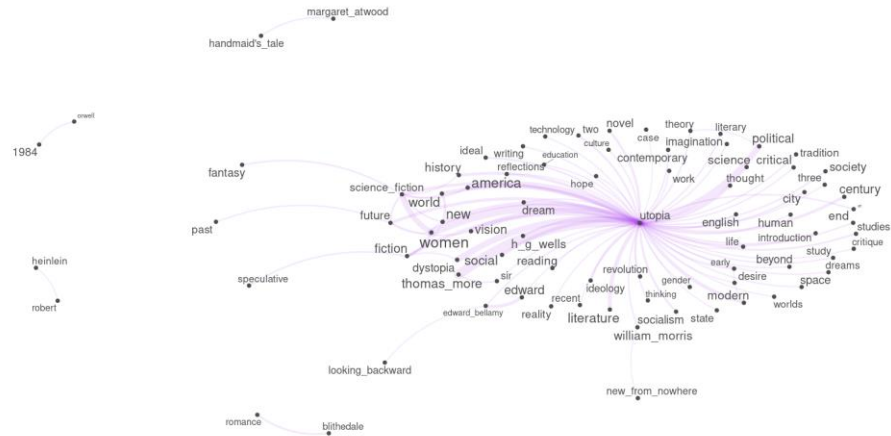


Fig. 2. Network analysis of English subset

The Feminist Subset is comprised of 453 works with about 358 different authors. When running the subset, we started by analyzing the number of words per title. We found that there are an average of seven words per title, the shortest title has one word, and the longest is 23 (see **Fig. 3**).

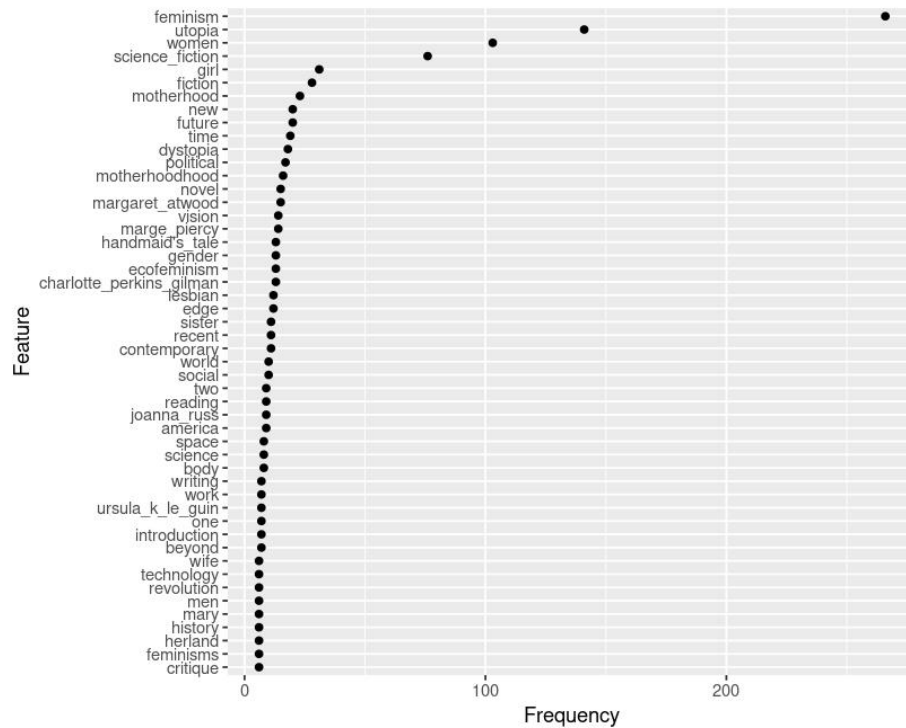


Fig. 3. Frequency of the words/themes feminist subset

We divided the distribution of papers per year into ten quantiles. We observed that between 1925 and 1982, we have 10% of the subset production, and between 1982 and 2022, the other 90% appear. We observed that the most productive periods are between 1982 and 1985 as well as between 1997 and 2000, while the least productive years are from 1925 to 1980, as well as from 2014 to 2022.

We then moved to the next phase, by obtaining the fifty most frequent words within the subset. As seen in Fig. 1, the most relevant words are feminism, utopia, and women, which are shown in titles over 100 times, as well as science fiction, which recurs slightly under 100 times.

In the top ten most frequent words, we can find keywords within feminism including 'feminism' itself, 'women', 'girl', and 'motherhood'.

In terms of the occurrence of author names, we observed that Margaret Atwood, Marge Piercy, Charlotte Perkins Gilman, Joanna Russ, and Ursula K. Le Guin are the most frequently mentioned in the subset, from most mentioned to least

mentioned. Consequently, since Margaret Atwood is the most frequently mentioned among the authors, *The Handmaid's Tale* shows up in eighteenth place, closely behind its author, as the only book title with enough expression to be considered in a token frequency table.

When conducting a network analysis with this subset we observed the emergence of four focus points such as ‘feminism’, ‘utopia’, ‘women’, and ‘science fiction’, which create an interconnected constellation.

Terms such as ‘new’, ‘future’, ‘science’, ‘contemporary’, and ‘recent’ are considerably connected to these four focus points. Author names such as Margaret Atwood, Marge Piercy, Charlotte Perkin Gilman, and Joanna Russ are also prevalent in the network, mostly related to the main constellation as well as their book titles (see **Fig. 4**).

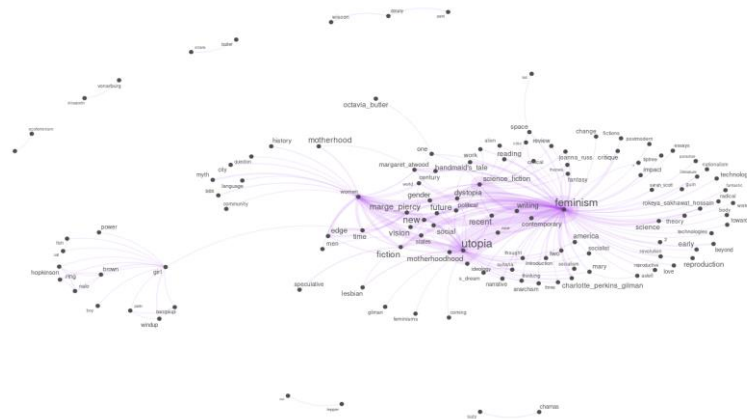


Fig. 4. Network analysis of the feminism subset

5. Discussion

The subsetting process thus leads us into a critical reflection of the results presented by the feminist subset, allowing us to offer insights into the construction of the field, potentially aiding the writing of critical feminist utopian literature in the future.

While terms related to ‘utopia,’ such as ‘utopian’ and ‘utopianism,’ are some of the most frequently used in the feminist dataset, we observed that dystopian novel

titles such as *Brave New World* (1932), *The Handmaid's Tale* (1985), and *Nineteen Eighty-Four* (1949) are the only ones to be represented in the token frequency table of the fifty most frequent words in the feminist subset, as well as in network analysis, which revealed the importance of feminist themes and their interplay with larger utopian and science fiction storylines, providing a broader perspective of the genre's evolution. Utopian titles, by contrast, are less visible. However, this discrepancy is not as apparent in the bibliography or literature review, suggesting an interesting disjunction between the focus of the feminist dataset and the actual representation of the metadata of utopian and dystopian works.

Contrary to the English subset database—which seemed to emphasize utopian ideals tied to themes like history and early modernism—the feminist subset's themes are more future-oriented, connected to terms such as 'future,' 'recent,' and 'contemporary.' This could suggest an emphasis on the forward-looking nature of feminist thought, imagining a progressive society that challenges past limitations. However, it is important to remember that our methods for identifying these themes rely on specific data visualizations and tools that can obscure their interpretative underpinnings. As Drucker (2011) reminds us, data visualizations—such as network graphs or frequency tables—are not neutral mirrors of reality but constructed representations shaped by human decisions at every step.

In this context, the distinction between "data" and "capta" becomes particularly relevant. Data is often treated as something objective and given, while "capta," as Drucker explains, emphasizes the active process of taking or constructing information. In our analysis of the feminist subset, what we consider "data" is, in fact, capta—information that has been selected, categorized, and interpreted according to our specific research questions and digital tools. The prevalence of the metadata of dystopian works in our dataset, for instance, is not a simple reflection of literary history but an outcome of the choices embedded in our data collection and the visualizations we employed. This realization challenges any notion that our results are self-evident; instead, they are interpretive, shaped by the very structures and frameworks through which we approach the data.

As we have seen before, a critical analysis of the field of utopian studies argues that no single definition of utopia is entirely satisfactory and that any characterization of utopia must account for its diverse facets. Gregory Claeys (2012) argues that utopia is *not* 'an exclusively literary tradition', yet the trend seems to be for utopian studies to be greatly focused on its literary expression. In our subset related to feminist utopianism, the works are overwhelmingly focused on specific authors and works of fiction, rather than on the other two faces of utopianism proposed by Lyman Tower Sargent (1994).

Analyzing the quantile graph on the subset related to feminism, we can see that the most productive moment in feminist critical works on utopianism came between the years 1983 and 1985, with 10% of all works in the subset having been written in these three years. Until 1985, only 20% of all the works had been written. *The Handmaid's Tale* was published in 1985, and it is an especially relevant work in the subset, as we see that an 'island' is formed by the name of this work and its author in the token network. This may indicate that *The Handmaid's Tale* is not only a reflection of a historical period of great interest in feminist utopianism but that it was also a catalyst for mass interest in studying utopianism through a feminist lens.

6. Conclusion

Through the use of digital humanities tools to analyze the metadata from the ARUS database on feminist utopian literature, this paper demonstrates the major themes, patterns, and connections that define the field. The primary goal of the research was to examine the metadata and explore how feminist utopianism has evolved and how specific works have contributed to this process. By applying methods such as token frequency analysis, network analysis, and quantile graphing, we traced the development of feminist utopianism and the roles specific works and authors played in shaping the discourse. Our approach demonstrates how digital humanities can facilitate large-scale literary analysis that goes beyond traditional methods, allowing us to visualize complex relationships and uncover trends that would otherwise remain hidden.

Johanna Drucker's argument that data is never neutral but always a product of interpretive processes can be central to our approach. In line with her view of data as capta, the visualizations we generated do not merely reflect the metadata but actively shape our understanding of it. This study shows how digital humanities tools, when critically applied, can offer new ways of thinking about literary history—revealing how feminist utopianism has evolved in response to broader cultural forces and highlighting underexplored connections between texts, authors, and themes.

Our use of digital humanities techniques, including distant reading of metadata and network analysis, allowed us to visualize recurring motifs such as 'utopia' and 'science fiction,' as well as the central role of specific authors. These methods revealed intricate networks of influence, offering a more comprehensive picture of the field than would be possible through close reading alone. By leveraging the power of metadata and visualization, the research has demonstrated how digital repositories, like the ARUS database, can enable new insights into large bodies of metadata, providing a platform for future research.

However, this study is not without its limitations. By focusing on titles alone, we acknowledge that future research could extend to other metadata fields such as abstracts or keywords, and explore additional variables like author gender or publication context.

Overall, this paper emphasizes the potential of digital humanities to deepen our understanding of literary fields by revealing underlying structures, trends, and biases. The findings not only offer critical insights into the evolution of feminist utopian literature but also underscore the importance of digital methods in transforming how we engage with literary analysis at scale. As digital repositories continue to grow, so too does the potential for these tools to reshape how we study and interpret literary history.

Acknowledgements

Financed by national funds through FCT - Fundação para a Ciência e a Tecnologia, I.P., under the project ref. UIDP/04097/2020.

<https://doi.org/10.54499/UIDP/04097/2020>

References

- Berry, D. M. (2012). Introduction: Understanding the digital humanities. In *Understanding Digital Humanities* (pp. 1–20). doi:10.1057/97802303719341
- Callaway, E., Turner, J., Stone, H., & Halstrom, A. (2020). The Push and Pull of Digital Humanities: Topic Modeling the “What is digital humanities?” Genre. *Digital Humanities Quarterly*, 14(1)
- Claeys, G. (2012). "The Five Languages of Utopia: Their Respective Advantages and Deficiencies; With a Plea for Prioritising Social Realism". In A. Blaim, & L. Gruszevska-Blaim (Eds.), *Spectres of Utopia* (pp. 26-31). Peter Lang.
- Drucker, J. (2011). Humanities approaches to graphical display. *Digital Humanities Quarterly*, 5(1), 1-21.
- Fancourt, D. (2004). *Altered states: Feminist utopian literature* (Doctoral dissertation, University of Leicester).
- 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>

- Green, Jen & Sarah Lefanu, eds. *Despatches From the Frontiers of the Ffmlk Mind*. London: The Women's Press, 1985
- Greenspan, B (2016). *Are Digital Humanists Utopian?: From Debates in the digital humanities 2016* (pp.393-409). Minneapolis, MN ; London : University of Minnesota Press
- Kirschenbaum, Matthew. (2010). *What Is Digital Humanities and What's It Doing in English Departments?* ADE Bulletin. 55-61. 10.1632/ade.150.55.
- Levitas, R. (2013). *Utopia as Method* (1st ed.). Palgrave Macmillan London.
<https://doi.org/10.1057/9781137314253>
- Marketos, P. R. (1993). The evolution of feminist utopias.
- Moretti, F. (2007), *Graphs, Maps, Trees: Abstract Models for a Literary History* (London:Verso).
- Moretti, F. (2013). *Distant reading* . (London:Verso).
- Sargent, L. T. (1994). The Three Faces of Utopianism Revisited. *Utopian Studies*, 5(1), 1-37. <https://www.jstor.org/stable/20719246?origin=JSTOR-pdf>
- Sargisson, L. (1995). *Contemporary feminist utopianism*. Routledge.