



Christina Dittmann, B.A.

An Exploratory Data Analysis of Sylvia Beach's Shakespeare and Company: Current Research, Possibilities, and Constraints

Master Thesis Exposé

Research Seminar

Georg Vogeler, Univ.-Prof. Dr.phil. M.A.

Graz, July 2024

Table of Contents

1. Introduction	3
2. Research Question	3
3. Contemporary Writing and Current Research.....	3
4. Material and Method	7
5. Possible Criticism and Challenges	9
6. Outline	7
7. Timeline and Workflow.....	11
Bibliography	13

1. Introduction

In the aftermath of World War I, emigrant Sylvia Beach opened an American bookstore in Paris that would in its course change Modernist literature culture fundamentally. Sylvia Beach was one of many expatriates in Paris when the city became the hotbed of art, cinema, literature, and music.¹ Her bookstore, which also functioned as a lending library, quickly established itself as a meeting point for the Lost Generation and other avant-garde movements.² It was at this address that expatriate writers such as Ernest Hemingway, F. Scott Fitzgerald, and James Joyce met, exchanged ideas, and influenced each other's work. This study explores the literary and social interactions within this cultural space and analyzes the patterns of reading habits and social connectivity among its author-members.

2. Research Question

Despite the importance of Shakespeare and Company in the accounts of Modernist literature, there has been limited quantitative analysis of its operations and the actual interactions among its members. Given that many members were writers themselves, it is particularly interesting to explore what these authors read.

This study aims to address this gap by examining the reading habits and network relationships within this community to understand how these elements influenced the development of Modernist literature. Therefore, the research question is: Are there discernible patterns of reading habits among the author-members of Sylvia Beach's Shakespeare and Company?

3. Contemporary Writing and Current Research

In 1956, Sylvia Beach published her autobiographical book titled *Shakespeare and Company*.³ The book provides valuable insights into her experiences of running the bookstore and her exchanges with numerous literary figures of her time.

¹ Mary McAuliffe, *When Paris Sizzled: The 1920s Paris of Hemingway, Chanel, Cocteau, Cole Porter, Josephine Baker, and Their Friends* (Rowman & Littlefield, 2016), 1–2.

² Noël Riley Fitch, *Sylvia Beach and the Lost Generation: A History of Literary Paris in the Twenties and Thirties* (New York, NY: Norton, 1985), 44–47.

³ Sylvia Beach, *Shakespeare and Company* (Lincoln: University of Nebraska Press, 1991).

Ernest Hemingway's memoir, *A Moveable Feast* (1964),⁴ provides a detailed and intimate portrayal of his determining years in Paris, particularly his interactions with fellow expatriates. Hemingway recounts his experiences at Shakespeare and Company, emphasizing the bookstore's central role for the Lost Generation. He portrays Beach as a nurturing figure whose support was essential to the literary community. Hemingway's narratives about Gertrude Stein, F. Scott Fitzgerald, Ford Madox Ford, and Ezra Pound highlight the social dynamics and relationships within the expatriate literary scene. His reflections reveal both the collaborative and competitive aspects of these interactions, marked by both personal and artistic challenges.

Moreover, Malcolm Cowley's *Exile's Return* (1934)⁵ and Robert McAlmon and Kay Boyle's *Being Geniuses Together* (1968)⁶ document the experiences and contributions of American writers in Paris during the interwar period. Both works highlight the essential role of literary salons like Stein's salon and bookstores like Shakespeare and Company in fostering Modernist literature.

Diana Souhami's biographical works, such as *Gertrude and Alice* (1991)⁷ and *No Modernism Without Lesbians* (2020),⁸ explore the contributions of women and LGBTQ+ individuals in Modernist literature. Figures like Beach, Stein, and Natalie Barney are depicted as deeply influential in inter-war Paris. Daniel Katz,⁹ Donald Pizer,¹⁰ and Craig Monk¹¹ contribute to this discourse by analyzing the impact of American expatriates, particularly focusing on the international exchange of ideas that shaped the literary output of the Lost Generation.

⁴ Ernest Hemingway, *A Moveable Feast* (London: Arrow, 1994).

⁵ Malcolm Cowley, *Exile's Return: A Literary Odyssey of the 1920s* (New York: Penguin Books, 1976).

⁶ R. McAlmon and K. Boyle, *Being Geniuses Together, 1920-1930* (Doubleday, 1968), <https://books.google.at/books?id=6R5bAAAAMAAJ>.

⁷ Diana Souhami, *Gertrude and Alice* (London: Pandora, 1991).

⁸ Diana Souhami, *No Modernism without Lesbians* (London: Head of Zeus Ltd, 2020).

⁹ Daniel Katz, *American Modernism's Expatriate Scene: The Labour of Translation*, Edinburgh Studies in Transatlantic Literatures (Edinburgh: Edinburgh University Press, 2014).

¹⁰ Donald Pizer, *American Expatriate Writing and the Paris Moment: Modernism and Place*, Modernist Studies (Baton Rouge: Louisiana State University Press, 1996).

¹¹ Craig Monk, *Writing the Lost Generation: Expatriate Autobiography and American Modernism* (University of Iowa Press, 2010).

Nora Benedict's article "Buyers versus Borrowers" (2022)¹² provides insights into the economic and logistical challenges of running a literary business abroad, which she exemplifies through Beach's efforts to support Irish writer James Joyce.

In 2024, *Modernism/modernity* and the *Journal of Cultural Analytics* published a special issue, "The World of Shakespeare and Company."¹³ The issue includes several articles which analyze historical and biographical data and thus explore reading habits, gender correlations, and other cultural dynamics at Shakespeare and Company.

4. Material and Method

In 2014, the project *Mapping Expatriate Paris: The Shakespeare and Company Lending Library Project* (MEP) was launched by Joshua Kotin, Jesse D. McCarthy, and Clifford E. Wulfman.¹⁴ Hosted by the Center for Digital Humanities (CDH), it has been guaranteed long-term support since 2020.¹⁵ During its course, the project was renamed into the *Shakespeare and Company Project*.¹⁶ As of 2024, the project leadership transitioned to Joshua Kotin, with Rebecca Sutton Koeser as the technical lead.¹⁷

The project's aim is to digitally recreate the cultural and intellectual landscape of interwar Paris, using documents from Princeton's collection Sylvia Beach Papers. The project itself focuses on Beach's bookkeeping and business documents, such as lending library cards, logbooks, and address books. A small number of these lending library cards, and other documents are also held by other institutions, including the Joyce Collection at the University at Buffalo Libraries and the Carlton Lake Collection at the Harry Ransom Center at the University of Texas at Austin.¹⁸ This digitization effort offers a glimpse into the operational and social dynamics of one of Paris' most influential literary centers.

The transformation process from the original source documents to the database can be seen in Figure 1. The lending library cards and logbooks were transcribed using TEI (Text

¹² Nora Benedict, "Buyers versus Borrowers: A Look at the Finances of Shakespeare and Company," *Digital Scholarship in the Humanities*, December 26, 2022.

¹³ Joshua Kotin and Rebecca Sutton Koeser, "The World of Shakespeare and Company: An Introduction," *Journal of Cultural Analytics* 9, no. 2 (May 29, 2024), <https://doi.org/10.22148/001c.116905>.

¹⁴ Andrew Thacker, "Review: Shakespeare and Company Project," *Reviews in Digital Humanities* 2, no. 7 (2021), <https://doi.org/10.21428/3e88f64f.f9558cc1>.

¹⁵ "Shakespeare and Company Project," The Center for Digital Humanities at Princeton, accessed June 22, 2024, <https://cdh.princeton.edu/projects/shakespeare-and-company-project>.

¹⁶ Hereafter, the *Shakespeare and Company Project* will be referred to as the *Project*

¹⁷ The project website can be found here: <https://shakespeareandco.princeton.edu>.

¹⁸ Joshua Kotin and Rebecca Sutton Koeser, "Shakespeare and Company Project Data Sets," *Journal of Cultural Analytics* 7, no. 1 (2022): 5, <https://doi.org/10.22148/001c.32551>.

Encoding Initiative) standards. While the cards were fully transcribed, the logbooks were only partially transcribed. This led to three TEI documents: cards, logbooks, and personographies. It was also at this stage that additional information was added to the data, such as local identifiers, links to VIAF accounts, normalized versions of names, gender, and nationality.

Following the transcription, the project transitioned into a relational database, using the Django framework. The database schema was initially designed and refined through iterative adjustments. Python scripts were employed to extract data from the TEI files and import it into the database, containing information about members, books, and events. It was at this stage that information from the address books was added to the database. Once the data was stored in the relational database, it was exported to CSV files, which will be the format used in this master's thesis.

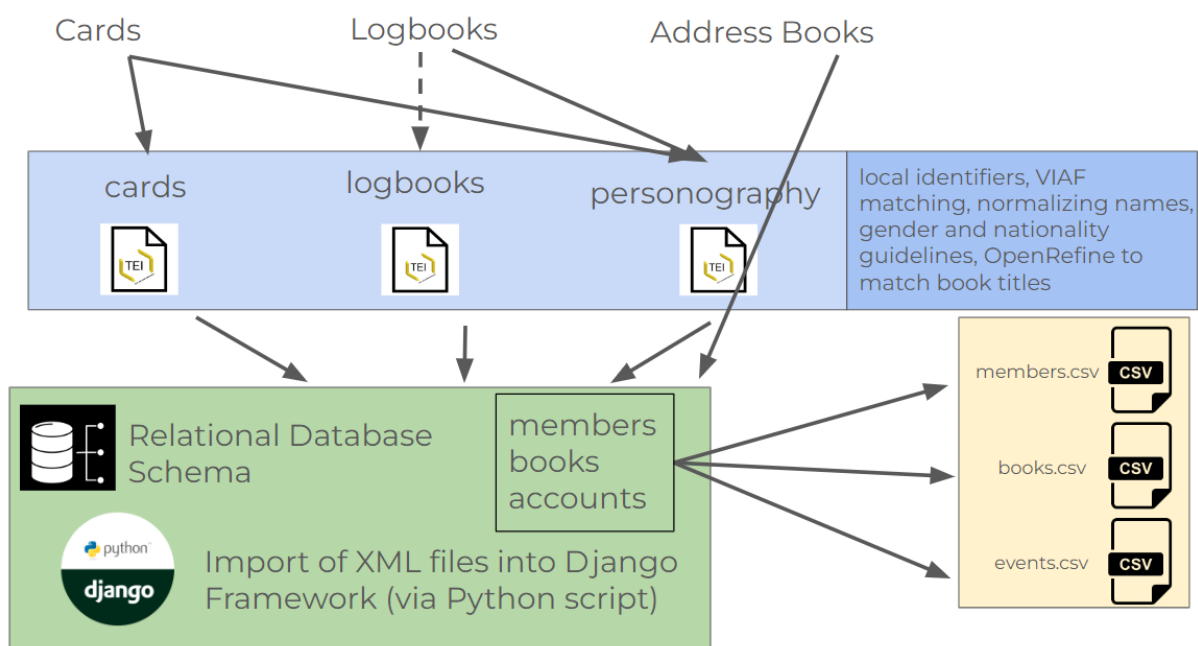


Figure 1, Transformation process

Based on the three datasets provided by the Princeton project, this study aims to implement **exploratory data analysis (EDA)** to investigate the aforementioned research question. The study uses various Python libraries for comprehensive analysis. Matplotlib is used for creating visualizations, while Seaborn enables improved customization of these visualizations. Pandas is employed for manipulating data frames, and NumPy is used for tasks such as entropy, binning, and other calculations. NetworkX is essential for network analysis, defining nodes (members) and edges (interactions like co-borrowing of

books), as well as calculating network diameter, identifying cliques, and detecting communities.

Data structuring is crucial for network analysis. Thus, different layout algorithms included in the NetworkX library arrange the nodes visually and highlight clusters of interactions. Metrics such as degree centrality and betweenness centrality measure each author-member's connectivity and influence within the network. Community detection algorithms identify groups within the network, indicating clusters of members with shared interests or literary tastes.

By integrating EDA and network analysis findings, this study provides a detailed understanding of the structural and substantive dynamics within Shakespeare and Company. Through careful methodological processes, both quantitative and qualitative data are combined to reveal complex patterns of literary exchange and social interaction, offering valuable insights into the role of Shakespeare and Company in the development of modernist literature.

5. Outline

The preliminary outline of the master's thesis is structured into two main parts: theory and analysis. The outline features only two levels for illustrative purposes:

-
- 1 Introduction
 - 2 Exploratory Data Analysis
 - 2.1 History and Developments
 - 2.2 Definition
 - 2.3 Goals
 - 2.4 Techniques and Procedures
 - 3 When Paris Sizzled and Modernism Shocked
 - 3.1 Modernism and the Obscure
 - 3.2 Dadaism, Surrealism, and the Celebration of Nothing
 - 3.3 Shakespeare and Company Opens its Doors
 - 3.4 Contemporary Writing and Current Research
 - 4 Princeton's Shakespeare and Company Project
 - 4.1 Sources

4.2	Transformation: From Historical Documents to A Relational Database
4.3	Datasets
4.4	“Le plan américain” of Bookkeeping
4.5	Limitations and Possibilities
5	Exploratory Data Analysis Workflow
5.1	Intersection of authors and members
5.2	Network Analysis
6	Findings
7	Conclusion
Bibliography	

The thesis begins with an “Introduction,” outlining the research objectives and their significance within the fields of digital humanities and American literature.

Following the introduction is the chapter “Exploratory Data Analysis” (EDA), which is divided into several smaller sections. “History and Developments” (2.1) provides a historical overview of EDA, while “Definition” (2.2) establishes a clear definition of the term, drawing on John Tukey’s works as well as modern explorations. “Goals” (2.3) details the primary objectives of EDA and highlights its separation from other forms of analysis. This subchapter is supported by “Techniques and Procedures” (2.4), which describes the methodologies employed in the thesis.

The chapter “When Paris Sizzled and Modernism Shocked” explores the cultural and societal changes in the post-war period. It specifically focuses on the alienating effects of World War I and the emerging emigrant culture. Within this chapter, “Modernism and the Obscure” (3.1) examines the impact of Modernism and outlines its key characteristics, exemplified by excerpts of influential literature. “Dadaism, Surrealism, and the Celebration of Nothing” (3.2) discusses these avant-garde movements and their connections to Modernism. “Shakespeare and Company Opens its Doors” (3.3) provides a detailed history of the bookstore and its influence on the Modernist movement. “Contemporary Writing and Current Research” (3.4) delineates contemporary memoirs, including Beach’s and Hemingway’s chronicles, and current scholarship related to the topic.

The chapter “Princeton’s Shakespeare and Company Project” summarizes the project’s aims and summarizes the transformation of historical documents from the bookstore. It includes “Sources” (4.1), which describes Beach’s original documents acquired by Princeton. “Transformation: From Historical Documents to A Relational Database” (4.2) outlines the data digitization and transformation process, while “Datasets” (4.3) details the content of the resulting datasets. “‘Le plan américain’ of Bookkeeping” (4.4) discusses Beach’s organizational strategies and membership conditions, sourced on Beach’s own memoir and historical documents. “Limitations and Possibilities” (4.5) evaluates the dataset’s constraints and potential, which is essential for a robust analysis and interpretation.

The chapter “Exploratory Data Analysis Workflow” describes the analysis process step-by-step, including initial content exploration. “Intersection of Authors and Members” (5.1) investigates relationships within the literary community, focusing on the members whose books were available in the lending library. It thus describes the joining and matching of two datasets – members and books. The subsequent section, “Network Analysis” (5.2), applies network theory to map similarities and interactions, such as the borrowing of the same books, between author-members.

The “Findings” section summarizes the study’s outcomes, and the “Conclusion” chapter reflects on its implications and suggests directions for future research. The thesis concludes with a bibliography of all works cited.

To ensure reproducibility, the complete code for data processing and analysis will be thoroughly documented and uploaded to GitHub. Apart from explanatory comments in the code itself, a comprehensive README file will include explanations of the project and detailed instructions for replication.

6. Possible Criticism and Challenges

One of the primary challenges in this study arises from the nature of historical data. The datasets, like many historical records, were not created with the intention of future digital analysis and are thus incomplete.

In their article “Shakespeare and Company Project Data Sets,” Joshua Kotin and Rebecca Sutton Koeser describe the incompleteness of the sources and thus of the datasets. They

state that one of the two given address books stops recording names mid-way.¹⁹ Furthermore, the cards and the address books occasionally include mere fragments of dates, “a year and a month, or just a year or a month, or no date at all.”²⁰ There are also instances where the cards and address books only include fragments of addresses, “hotels without street addresses, home addresses without street numbers.”²¹ It must be noted, however, that by focusing on a smaller group, in this case author-members, the given data is more comprehensive.

Additionally, Koeser and LeBlanc argue in their article “Missing Data, Speculative Reading” that the purchase activity represented in the data is not as expressive as the borrowing activity.²² Beach recorded the purchase activity in separate documents. Only occasionally did she note purchases in a member’s card.²³ As the borrowing and purchase data derive solely from the cards, the borrowing data offers a more consistent and reliable source for analysis.

It is crucial to acknowledge the gaps and inconsistencies within the data. Despite these inherent limitations, the dataset remains rich in potential. By focusing the analysis on a specific group, such as the author-members, it is possible to minimize the impact of these gaps. Analyzing common borrowing behavior among author-members provides a reliable foundation, allowing for valid pattern identification. Even though additional similarities might have been observable if the data were complete, the existing data enables interpretation of given linkages. For example, if the data shows that both Hemingway and Joyce borrowed the same five works by William Faulkner, while Hemingway and Stein borrowed completely different books, we can ascertain definite linkages between Hemingway and Joyce. It is plausible that further linkages between Hemingway and Stein might emerge if the data were complete. Nevertheless, the given data enables researchers to identify and analyze the minimum set of linkages between reading habits. This approach ensures that the analysis remains as robust as possible.

¹⁹ Kotin and Koeser, 10.

²⁰ Kotin and Koeser, 10.

²¹ Kotin and Koeser, 10.

²² Rebecca Sutton Koeser and Zoe LeBlanc, “Missing Data, Speculative Reading,” *Journal of Cultural Analytics* 9, no. 2 (May 29, 2024): 4–5, <https://doi.org/10.22148/001c.116926>.

²³ Koeser and LeBlanc, 4–5.

7. Timeline and Workflow

As of July 8, 2024, I am at the midpoint of my research and writing process. The initial phase of brainstorming and selecting a topic was completed by December 31, 2023. This phase involved thorough research and consultation with Martina Scholger to finalize a suitable and impactful research topic.

From January 1 to February 28, 2024, I conducted an extensive literature review, which provided a theoretical foundation for my study. Subsequently, from March 1 to April 30, 2024, I focused on understanding the data, which included preliminary analysis and grasping the complex digitization and transformation process.

I am currently engaged in writing the theoretical framework and performing EDA, both scheduled to conclude by July 15, 2024. The theoretical writing involves defining key concepts and hypotheses (related to both EDA and Modernism), while the practical implementation of EDA uncovers patterns in the data through statistical techniques. This practical part has led to my narrowing down the research question to author-members and opting for network graphs to visualize author-member connections.

Following this, I will commence writing up my findings from July 1 to August 30, 2024. This stage involves documenting results, interpreting them, and discussing their implications. The documentation phase will take place from September 1 to September 30, 2024, where I will compile all sections of the thesis and ensure consistency as well as coherence. It is until this stage that I will allow myself to alter and improve my code, depending on feedback from Martina Scholger and fellow students. It is vital for me to follow the time schedule while remaining enough flexibility and dynamicity to further align the work with in-coming feedback. The final phase is proofreading, scheduled from October 1 to November 15, 2024. This step involves reviewing the thesis to correct errors and enhance clarity. The timeline, graphically represented in Figure 2, ensures systematic progress as well as flexibility.

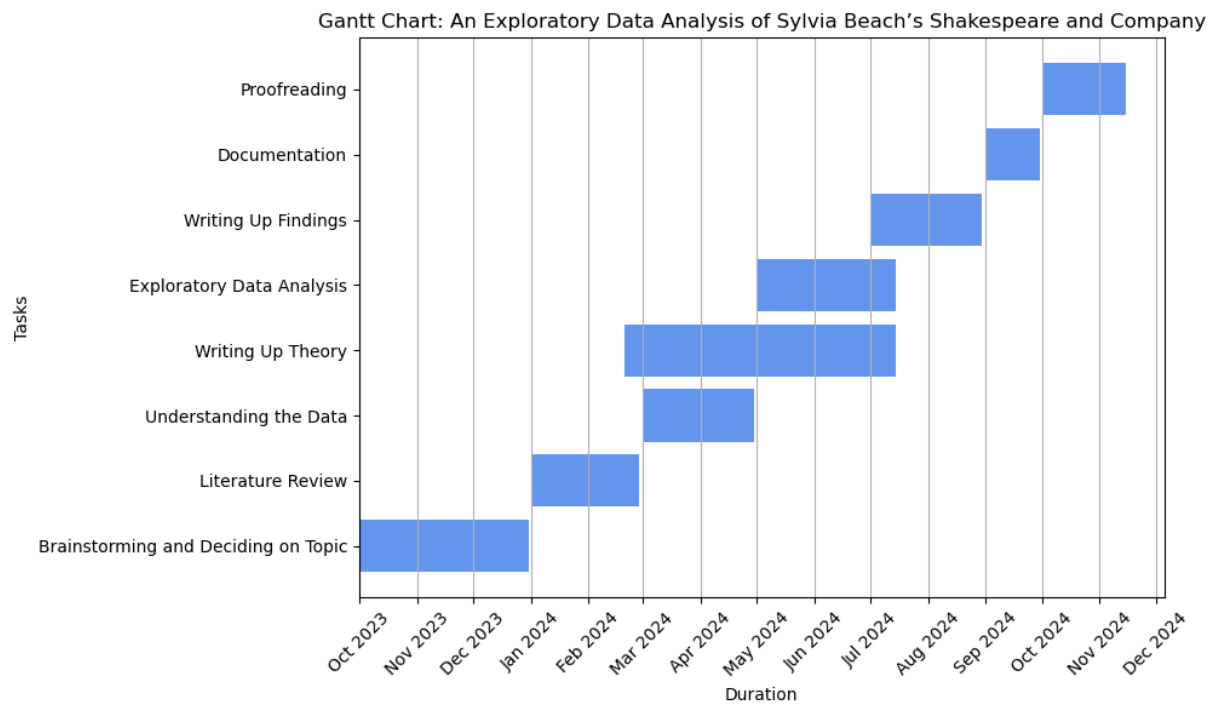


Figure 2, Gantt Chart, created with Python. Code can be viewed here: <https://github.com/dittm/gantt-chart.git>

Bibliography

- Beach, Sylvia. *Shakespeare and Company*. Lincoln: University of Nebraska Press, 1991.
- Benedict, Nora. "Buyers versus Borrowers: A Look at the Finances of Shakespeare and Company." *Digital Scholarship in the Humanities*, December 26, 2022.
- Cowley, Malcolm. *Exile's Return: A Literary Odyssey of the 1920s*. New York: Penguin Books, 1976.
- Fitch, Noël Riley. *Sylvia Beach and the Lost Generation: A History of Literary Paris in the Twenties and Thirties*. New York, NY: Norton, 1985.
- Hemingway, Ernest. *A Moveable Feast*. London: Arrow, 1994.
- Katz, Daniel. *American Modernism's Expatriate Scene: The Labour of Translation*. Edinburgh Studies in Transatlantic Literatures. Edinburgh: Edinburgh University Press, 2014.
- Koeser, Rebecca Sutton, and Zoe LeBlanc. "Missing Data, Speculative Reading." *Journal of Cultural Analytics* 9, no. 2 (May 29, 2024). <https://doi.org/10.22148/001c.116926>.
- Kotin, Joshua, and Rebecca Sutton Koeser. "Shakespeare and Company Project Data Sets." *Journal of Cultural Analytics* 7, no. 1 (2022). <https://doi.org/10.22148/001c.32551>.
- Kotin, Joshua, and Rebecca Sutton Koeser. "The World of Shakespeare and Company: An Introduction." *Journal of Cultural Analytics* 9, no. 2 (May 29, 2024). <https://doi.org/10.22148/001c.116905>.
- McAlmon, R., and K. Boyle. *Being Geniuses Together, 1920-1930*. Doubleday, 1968. <https://books.google.at/books?id=6R5bAAAAMAAJ>.
- McAuliffe, Mary. *When Paris Sizzled: The 1920s Paris of Hemingway, Chanel, Cocteau, Cole Porter, Josephine Baker, and Their Friends*. Rowman & Littlefield, 2016.
- Monk, Craig. *Writing the Lost Generation: Expatriate Autobiography and American Modernism*. University of Iowa Press, 2010.
- Pizer, Donald. *American Expatriate Writing and the Paris Moment: Modernism and Place*. Modernist Studies. Baton Rouge: Louisiana State University Press, 1996.
- Souhami, Diana. *Gertrude and Alice*. London: Pandora, 1991.
- . *No Modernism without Lesbians*. London: Head of Zeus Ltd, 2020.
- Thacker, Andrew. "Review: Shakespeare and Company Project." *Reviews in Digital Humanities* 2, no. 7 (2021). <https://doi.org/10.21428/3e88f64f.f9558cc1>.
- The Center for Digital Humanities at Princeton. "Shakespeare and Company Project." Accessed June 22, 2024. <https://cdh.princeton.edu/projects/shakespeare-and-company-project>.