Ca' Foscari
University
of Venice

Master's Degree Programme

In Digital and Public Humanities

Final Thesis

Analysing and Reaccessing texts through Character Interaction Network

The Interpoetic Narrative of Spoon River Anthology

Supervisor
Prof. Fabiana Zollo

Assistant Supervisor
Prof. Franz Fischer

Graduand
Mattia Moscato
Matriculation number 881864

Academic Year

2023/2024

TABLE OF CONTENTS

ABSTRACT

This study employs Character Network Analysis (CNA) to examine Edgar Lee Masters' Spoon River Anthology, a distinctive collection of interconnected poems. It addresses the limited application of CNA to unconventional literary works by evaluating its effectiveness in this context. The research focuses on how character interactions construct an interpoetic narrative within the anthology. By integrating manual data extraction, network visualisation, and sentiment analysis, the study maps and interprets these interactions. Findings reveal a central connected component with key characters arranged in "Hub-and-Spoke" structures that drive the narrative. The analysis also considers the impact of supplementary poems from the 1916 edition, demonstrating that these additions reinforce but do not significantly alter the existing narrative. Furthermore, the study highlights how poems often express negative sentiments towards those poems/characters inspired by real Lewistown residents, reflecting Masters' critical view of his hometown. By employing CNA, the research provides a novel approach for re-accessing and re-evaluating the text, showcasing CNA's potential in literary analysis. The findings suggest that CNA could be a valuable tool for future research in digital humanities, advocating for the creation of a comprehensive digital edition to enhance accessibility and engagement with the text.

INTRODUCTION

"Once upon a time" may not be the most fitting way to introduce a master's thesis, but it is an appropriate preamble for something as intricate as the domain in which this research is situated. Indeed, we find ourselves within the interdisciplinary field of Digital Humanities, which extends beyond merely applying computational methods to traditional humanities research, which are gradually embracing the arrival of these new methods to assist them. Digital Humanities holds a distinctive place in academia, where its scope may vary from being an independent discipline to a research field or study area. This variability highlights the field's adaptable boundaries and its ability to engage with and transform existing disciplines (*Sahle, 2015*). The foundational principles of Digital Humanities have been articulated since its early development. For example, the manifesto published in 2011, following discussions at THATCamp in 2010, outlines core values such as open data access, the free exchange of methods, and the integration of digital tools into educational practices (*Dacos, 2011*). These values support collaboration, innovation, and the development of scalable digital infrastructures that utilize collective expertise to establish best practices across disciplines. This research reflects these values, recognizing Digital Humanities as a complement to traditional methods, highlighting its value through quantitative analysis while maintaining the importance of qualitative analysis (*Krämer, 2023, p. 17*). Eventually, there must be a consensus among scholars that theories built on one-sided instruments have limited chances against the tools that the Digital Humanities are currently adopting (*Windhager & Mayr, 2024, p. 25*). Investigating even further, this research specifically aligns with digital literary criticism, focusing on narratology and employing Character Network Analysis for narrative analysis. Digital literary criticism, like most components of the Digital Humanities, initially experienced a phase in which there was a tendency to apply new tools to an old medium, often basing the majority of new possibilities on this approach while still allowing space for qualitative research (*Murray, 2015, pp. 317-318*). It now finally relies more on the use of computational methods like text mining and Natural Language Processing (NLP) for a more automated literary analysis. Among these methods, network analysis has also proven useful (*Sharma et al., 2024, p. 5235*). Given the nature of the field, it relies heavily on the digitization of literary works for analysis and the creation of datasets. However, due to the challenges faced by non-digital humanities scholars, there is a growing trend towards providing

scholars with digital tools through information visualisation instead (*Trocchianesi & Bollini, 2023*). One area explored by digital literary analysis is narratology, which has seen renewed interest due to automated methods for extracting and analysing textual features (*Nelis, 2017*). Nevertheless, narratology is a field with a long tradition, as studying narrative and the act of narrating is one of the defining characteristics of human experience, even though its formal theorization solidified in the 1960s (*Prince, 1991 ,p. 533*). It must pay close attention to various aspects, such as story and its narration, often blending theorization with the practical study of narrative in search of its structures and optimal compositions, fostering a solid and amicable relationship between the work and its audience. Given its practices, a new perspective has been necessary to shift the focus of this subject toward specific digital methods, namely the construction of narrative structure as primarily interactions between characters, thereby addressing the issue of agency (*Franzosi et al., 2012*). Character Network Analysis is one such method that is well suited for this new perspective on narrative and is currently used to study narrative structure (*Rochat, 2014*). This thesis will offer an overview of narratology and its relationship to the application of Character Network Analysis, which will be discussed in the first chapter. Moving from the general to the specific, this thesis will review the current framework of the methodology, seeking gaps in the existing literature and proposing a project aimed at exploring them. In particular, it will address the lack of reuse of these networks for text analysis and the potential to engage with literary works beyond conventional narrative texts. Information visualisation remains an area that is not widely embraced by humanists, who primarily focus on the digitization of works in digital scholarly editions, where data visualisation could greatly enhance access to texts. To facilitate this meta-research, Edgar Lee Masters' *Spoon River Anthology* has been chosen for various reasons, including my personal interest, as the work is connected to the history of Italian publishing and singer-songwriting. The work's structure allows for a perfect application to study both character interactions and possibilities for re-access. It also aims to benefit humanities scholars due to its limited success in digital and academic contexts. The second chapter will introduce the history of the work and its main features, focusing on aspects that facilitate narrative analysis. The analysis will follow the IMRAD structure, with a section dedicated to outlining the methodologies, one presenting and analysing the results, and another discussing the findings. Reaccess to the work through Character Network Analysis will be simulated by imitating the workflow of a humanist without

digital skills and one with such skills. For the former, close reading will be employed through the structures studied in the interaction network, while the latter will apply Sentiment Analysis. Consequently, the discussion will directly incorporate these subsequent analyses. These structural and methodological considerations are not arbitrary. They are, in fact, central to the Digital Humanities and contribute to the ongoing debate regarding the value of this interdisciplinary field (*Joyeux-Prunel, 2024*). The conclusions will therefore include both a summary of the case study and responses to the initial questions, thus closing this ring composition. Final efforts were made to align the project with values of collaboration and dissemination, which are often overlooked by humanists (*Ruediger & MacDougall, 2023*). This includes drawing conclusions from main findings and considering ways to improve and share research findings, reflections on methodology, discoveries about the work, and its digital future. As part of these objectives, the creation of a website for the thesis project and a repository for all collected and used materials are included.

1. FROM NARRATOLOGY TO CHARACTER NETWORK ANALYSIS

Interest in narrative studies has long been present in academic circles, experiencing a steady increase often referred to as the "Narrative Turn". This growth can be linked to three primary factors: first, there is an epistemological motivation, recognizing narrative as a fundamental aspect of human experience that helps individuals interpret their surroundings; second, there is a historical-philosophical connection to Postmodernist themes, leading to the literarization of theory and the theorization of literature. Finally, there is a cross-disciplinary need for a structured theoretical framework to describe narrative phenomena, roles, and effects, using categories and models from narratology (*Sosič, 2022, pp. 6-7*). For instance, storytelling practices are applied in fields such as marketing and management, extending the area of study not only to textual contexts but also to mixed text and image ones, such as online social networks (*König, 2022*). Over the past decade, narratology has gained prominence, largely due to advancements in Natural Language Processing. These technologies have expedited research processes, particularly in the humanities (*Piper et al., 2021, p. 298*). Consequently, scholars have started exploring new methods for studying narrative, addressing areas previously overlooked due to the limitations of quantitative approaches in literary studies (*Chen et al., 2019, p. 800*). These new methods are still rooted in the concepts and areas of study and analysis within the field of narratology. Although narratology originally emerged within the academic field primarily for literary studies, it has become increasingly complex and filled with nuances, especially in recent times. Starting from its own terminology, in fact, the terms "Narratology" and "Narrative Theory" are often used interchangeably (as can be seen from: *Amerian & Jofi, 2015*) or to describe distinct fields. Narratology primarily examines the structural and formal aspects of narratives, while narrative theory includes interpretive and contextual dimensions, analysing the role of narratives in society and culture (*Sosič, 2022, pp. 7-9*). This divergence appears to stem from a shift in focus from literary studies to sociology, where narrative began to be studied as a key element of human social identity. Despite this shift, both Narrative Theory and Narratology engage in narrative analysis, defined as research on and with narratives (*Wang et al., 2020, pp. 3-4*). The challenge remains as the definition of narrative varies by field, with no universally accepted standard for identifying a sequence of elements as a narrative or defining the concept itself (*Piper et al., 2021, p. 298*; *Santana et al., 2023, pp. 8395-8396*). In narratology,

narrative refers to how a story is told or interpreted, consisting of the role of characters and events, and their causal or temporal relationships, which together form the plot (*Min & Park, 2019, pp. 1*-2). The various elements that constitute it are those considered when conducting a narrative analysis. The distinction between telling and perceiving a narrative its focal to narratology, in fact it divides classical from post-classical studies. Classical studies focus on narrative production and its interpretation, while post-classical studies consider reading and reception (*Sosič, 2022, p. 10*). Thus it would appear that postclassical studies primarily focus only on the significance of the reader and the audience, drawing substantial inspiration from Narrative Theory. But despite this significant difference, which pertains to the shift in postmodernism towards a study of reception rather than the pursuit of formal values, classical and postclassical studies seem to both influence one another and draw upon each other's findings. Central to both is the concept of narrative structure, which describes how a narrative is constructed, developed and engages the reader by evoking deep emotions (*Berhe et al., 2023, p. 64*). This relies in the clear distinction between how a story is presented and the story itself (*Cohn, 2013, p. 416*). This distinction is referred to as the difference between "story" and "discourse," and it is crucial, as analysing both elements allows us to define a genre or determine what a certain type of work should include to be classified in a particular way (*Chatman, 1975, pp.295-296*). This is necessary because it assumes that the same story can be told differently, while with the same discourse can convey different stories. Therefore, various methodologies, depending on the period and the specific literary works in question, have sought to determine whether structural narrative patterns exist or how a story should be told, which is essentially what is meant by studying on narrative and with narrative. These two have influenced each other and continue to do so. Consider, for instance, that the 2013 film *Locke*, shown out of competition at the 70th Venice Film Festival, shows a composition probably inspired by Aristotle's three unities of dramatic composition: unity of place, time, and action (*Krijanskaia, 2008, p.338*). Hence accompanying this theorisation of how a narrative should be constructed is the study of the structure itself, both from a discursive perspective and in relation to the story, as well as the patterns identified within works. This approach is more aligned with classical narratology and structuralist studies, such as those by Todorov. This is a very interesting aspect of the study of narrative, namely that it appears that theorization, study and story writing go hand in hand, and it is possible to examine the most well-known theories to observe how stories have primarily been

recounted and narrated. One of the most well-known is Freytag's pyramid, a theory that divides a story into the exposition of a problem and its final resolution, with a central climax that is gradually approached and then descended from, as the story progresses from beginning to end (*Ciğerci & Yıldırım, 2024, pp. 5701-5702*). It underscores the previously mentioned relationship between theory and study in this process, even though it may not seem apparent. Freytag, in fact, develops this theory based on classical Greek tragedy, a form of content that is now studied as literary material but was primarily created to be performed in a theatre (*Francis, 2020, p. 13*). For this reason, as a theory for constructing narratives, it refers more to a visual and auditory component rather than a literary one and suggesting that those who applied this theory to their narrative construction created a work rooted in the tradition of tragedy. These considerations are also reflected in more recent essays that discuss how writing and analysing a narrative seem like two different actions, yet both inevitably draw from similar abstract models (e.g., *Gervás et al., 2006, p. 44*). This example further illustrates that a narrative cannot be effectively constructed, nor can its mode of narration be selected, without first accounting for the medium through which it will be received. When imagining Freytag's pyramid, its visual representation helps us understand that a story must also be constructed and narrated in a "spatial" sense, meaning that the climax should occur midway through the narrative as can be seen from Figure 1.

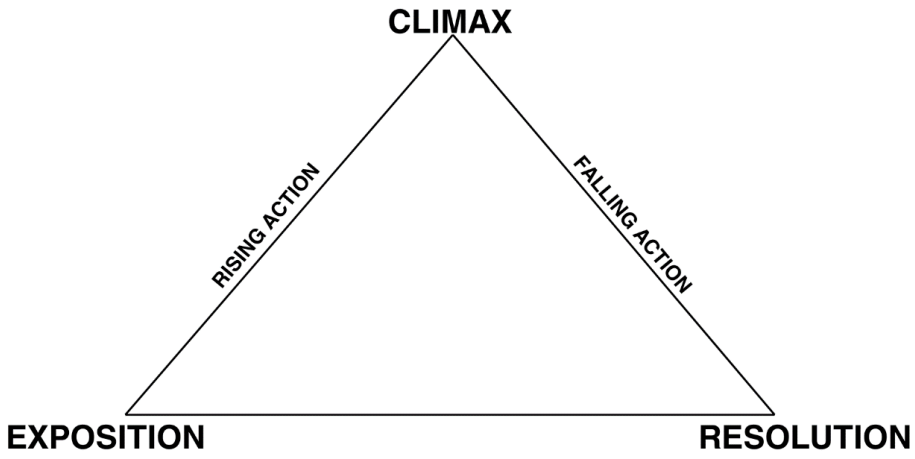

Fig. 1: Visualisation of the pyramid based on Freytag's theories

Appropriate adjustments should then be made, considering the medium used to tell the story. For example, in a book, the climax would typically be positioned around the middle, and similarly, in a film, it would occur around the midpoint of the viewing experience. Yet, it's also essential to consider who is engaging with the narrative and how they are experiencing it in these different formats so the difference between reading and watching. In my view, this is what connects classical and postclassical studies: it seems impossible to conceive of a narrative without also considering the audience to whom it is told and the medium portraying it for its analysis. These studies, therefore, vary by focusing on different aspects of this relationship according to their specific areas of expertise or scholarly traditions. This preamble opens the necessities and challenges of studying narrative through quantitative methods. Since the theorization of narrative is closely tied to the effort of visualizing narrative structure as well as describing it, it is reasonable to assert that quantitative analysis was not only helpful but that relying on data visualisation significantly improved comprehension. For functioning, quantitative narrative analysis mostly employs invariant linguistic properties of the text itself (*Franzosi, 2010, p. 51*). This action is grounded in the complexity of natural texts and the various ways to define their lexical and sublexical attributes, which subsequently contribute to the formation of sentences and phrases (*Xue et al., 2019 p. 2*). This stage is often referred to as text segmentation and prepares for subsequent analysis (*Beeferman et al., 1999 , pp. 177-178*). This method facilitated straightforward operations for narrative study, thanks to Natural Language Processing algorithms that would automate this process this process to allow for in-depth and automated study. (*Li et al., 2018, pp. 4166*). This shift is attributed to how new technologies have progressively enhanced text interpretation, enabling a focus not only on purely formal linguistic aspects but also on the narrative as a direct connection to the agency of the characters involved, thereby shaping the various events (*Tebaldi et al., 2019 p. 3069*). Training artificial intelligence to perform those kinds of analyses following segmentation entails several works. This concerns the specific tasks that the artificial intelligence must perform: identifying characters or agents and interpreting the relationships between them within the context; grasping the concept of temporality to determine whether events occur in a sequence or across multiple timelines; recognizing the setting to distinguish activities that belong within or outside that environment; and evaluating whether the narrative is presented from a specific perspective, differentiating between subjective and objective viewpoints. Each of these tasks requires targeted

training methods (*Piper et al., 2021, pp. 299-304).* An example of this is Sentiment Analysis, a methodology that plays a significant role in narratological analysis. With this approach is possible to examine the narrative perspective and sequence, understanding subjective from objective sentences (*Mukherjee & Bhattacharyya, 2013, pp. 22-23*). This is made possible through machine learning (ML) and deep learning (DL) algorithms, which are applied according to specific needs. Machine learning requires a dataset with labelled data and features, whereas deep learning algorithms, including advanced Large Language Models like GPT, DialogueLLM, and InstructERC, are employed to identify and interpret textual patterns (*Jim et al., 2024 , p. 11-16*). The use of these new technologies enables an analysis that, based on the emotions expressed by a character or the words in the examined segments, can be valuable for visualizing the narrative structure during moments of calm or conflict between characters, as well as identifying peaks of sentiment in specific sections of the text. This directly offers a more or less accurate representation of a narrative structure, but it is entirely dependent on the ability of algorithms to correctly classify and contextualize sentiment. Among the quantitative methods for the study of narratology, in line with the wave of reconsidering narrative from the perspective of agency, there is an analytical methodology that is specifically suited on it and aims to reconstruct it from character interactions using graph theory: Character Network Analysis. Unlike other quantitative approaches that analyse narrative arcs or structures through text segmentation and functional or cognitive words (e.g., *Boyd et al., 2020*), Character Network Analysis, also known as CNA, directly examines how characters interact with each other. Characters, as nominal entities, and their actions (events) are crucial to the story's development and its central component (the plot), making them key subjects of narrative analysis (*Srinivasan & Power, 2022, pp. 20-21*). It should be noted, however, that this type of analysis sidelines one of the central roles of narration, specifically the narrator, instead seeking to focus primarily on the development of main or secondary events based on the actions and interactions themselves. Derived from Network Analysis, a well-established method in Digital Humanities (*Ahnert et al., 2020*), Character Network Analysis uses a network graph where nodes represent characters and edges illustrate their interactions *(Lee & Jason, 2020, pp. 2-3*). Initially gained attention for its effectiveness in revealing narrative structures (e.g., *Min & Park, 2016*), classifying literary genres (e.g., *Ardanuy & Sporleder, 2015*), quantitatively confirming literary theories (e.g., *Elson et al., 2010*), and identifying character roles (*Kwon & Shim, 2017*). This

was made possible by the paradigm shift towards the study of agency and the ability to use Network Analysis and its visualisation to study the dynamics between characters, much like the analysis of an actual interconnected social system (*Agarwal et al., 2021, p. 217*). In fact the main activities and potential of the methodology concern theories originating from the field of sociology. For example, it reveals that similar literary genres contain shared interaction patterns, and that metrics typically used in social network analysis can highlight character roles, such as the protagonist or antagonist. However, it was previously considered labour-intensive due to the manual data extraction, as early technology struggled to accurately recognise and extract data about characters and their interactions (*Bimantara et al., 2024, p. 3*). This occurs because a single character may be referred to by different names or pseudonyms, requiring contextual reading to accurately identify who is being referenced. Similarly, the types of interactions being referenced can vary widely, including actions or conversations, and can be presented in a more explicit or implicit manner. A significant advancement came with the 2019 publication of *Extraction and Analysis of Fictional Character Networks: A Survey* by Labatut and Bost, which introduced a more standardized methodology and brought together the majority of projects that employed this approach, making it one of the key papers outlining the state of the art in CNA (*Labatut & Bost, 2019)*. The authors also identified and exposed inefficiencies in extracting characters and interactions from texts, thereby spurring interest in methods that employ Natural Language Processing to automate data extraction for this type of analysis.

1. 1 Framework for Character Network Analysis

Thanks to the work by Labatut and Bost in 2019, it is possible to have a complete vision of the steps that comprise the initial phase of CNA, namely the preparation of the interaction network, which is essentially the process of extracting the character network itself from the work of fiction under study, as shown in Figure 2, since this methodology is, of course, not only applied to texts but also to television series, films, and other narrative forms.

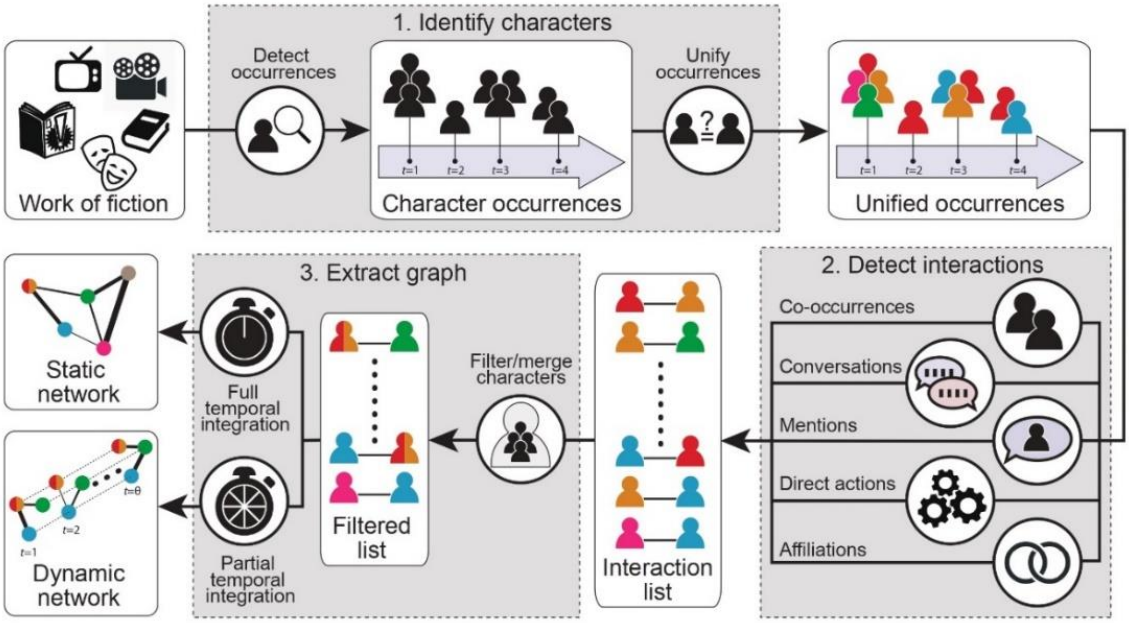

Fig. 2: Overview of the generic character network extraction process framework created by Labatut and Bost

The first and second phases described by the framework, identifying characters and detecting interactions (such as co-occurrences, conversations, mentions, direct actions, and affiliations), are typically the most time-intensive. The third phase, which closely resembles a data modeling phase, focuses on extracting the graph based on the annotated attributes and deciding whether to create a static or dynamic network based on the temporal value of the interactions. While AI systems trained to handle character identification and interaction detection are now more and more available, many scholars in the humanities still lack the digital skills needed to automate this process or train AI specifically for their chosen work of fiction. This often leads to a manual process of data extraction and annotation, which demands extensive reading and repeated verification of the dataset (*Labatut & Bost, 2019, pp. 6–21*). Even when using natural language processing tools, full automation is uncommon, as names, nicknames, and pseudonyms of characters frequently need to be manually entered into the dataset (e.g., *Fan & Li, 2022, p. 2*). The most recent tools are beginning to automate this process, including Renard (Relationships Extraction from NARrative Documents), a modular pipeline capable of extracting both static and dynamic networks from a narrative text, and available as a Python library designed to be accessible for humanists (*Amalvy et al., 2024*). Subsequent analyses are mostly object-oriented, meaning they are based on the initial research questions, which are currently divided according to the findings already discussed: confirmation of literary

theories, degree of similarity to reality or fiction, role detection, narrative structure, and story decomposition. When confirming literary theories, the research is organized specifically around the selected theory, exploring potential correlations. In contrast, when examining the level of similarity between the work of fiction and reality, a comparative method is employed. This refers to the need to compare the character network with real-world data using similar metrics, or by employing visualisations specifically designed for this type of analysis and adjusting the comparison parameters according to the research objective (*John et al., 2019, pp. 248-249*). Role detection is more standardised and primarily based on centrality measures, which are topological in nature (*Labatut & Bost, 2019, p. 28*). Story decomposition, and consequently narrative structure, often rely on mixed methods applied to dynamic networks, such as combining topological analysis and network analysis measures with sentiment analysis (*Min & Park, 2016, pp. 1-2*). Despite the methodology appearing solid at first glance given the results achieved thus far, it presents several challenges. One significant issue is the lack of suitable ground truth datasets specifically designed for fictional narratives. Most existing datasets focus on non-fiction, which creates a notable gap that raises important questions regarding the reliability of various extraction techniques. Furthermore, the tools used for character network analysis often follow traditional methodologies rather than seeking more specific and suitable approaches for the type of analysis. In conclusion, while effective in some contexts, these approaches frequently overlook the complexities of narrative structures (*Labatut & Bost, 2019, p. 30-34)*. However, as is common across various domains within the Digital Humanities, substantial collaboration could help address this issue and facilitate effective data interpretation, considering that narratology is a multifaceted field with a rich tradition (*Piper, 2023, p. 35*).

1.2 Contribution of the thesis

Considering the current state of Character Network Analysis, my research aims to align its framework with the values of Digital Humanities discussed in the introduction. This involves not only conducting research that aids in understanding this methodology and its applications, but also addressing the needs of scholars working with the designated case study. As previously discussed, narratology is a complex field, as are the research methodologies that employ criteria closer to scientific standards. This necessitates

envisioning the role of the Digital Humanist as a bridge between qualitative research and its traditions, and quantitative research. Researchers are continually improving methods for accessing, visualizing, and studying textual data tailored to various types of literary works or collections. This has prompted discussions on structuring collaboration between humanists and designers (*Trocchianesi & Elli, 2023, pp. 26-27*) and on how information and data visualisation can enhance literary criticism, especially considering the differences between qualitative and quantitative research (*Elli, 2023, pp. 159-161*). In this collaborative context, scholars are not only applying CNA but also exploring ways to provide non-digital humanists with these tools (*Chen et al., 2023*). Typically, Digital Humanities would opt for accessing digital textual information using methods similar to those employed for print materials, such as studying methods to produce digital editions that incorporate multiple versions, editions, and digital tools. A workshop organized by the *Institut für Dokumentologie und Editorik* in March 2022 led to a manifesto that addressed the creation of scholarly digital editions, focusing on methodological, practical, and organizational dimensions (*Ciotti et al., 2022, p. 103*). Primarily, however, these approaches aim to leverage the benefits of new technologies without straying too far from tradition. Data visualisation in both quantitative and qualitative contexts can serve as an excellent method for reevaluating texts and assisting digital literary criticism, such as the project *Atlante Calvino*, which enables the visualisation of the author's work through data representations of the texts and recurring themes (*Falcetto, 2019*). The atlas allows for the exploration of works through an interface featuring a thematic path and various dynamic visualisations based on the selected theme, as illustrated in Figure 3.

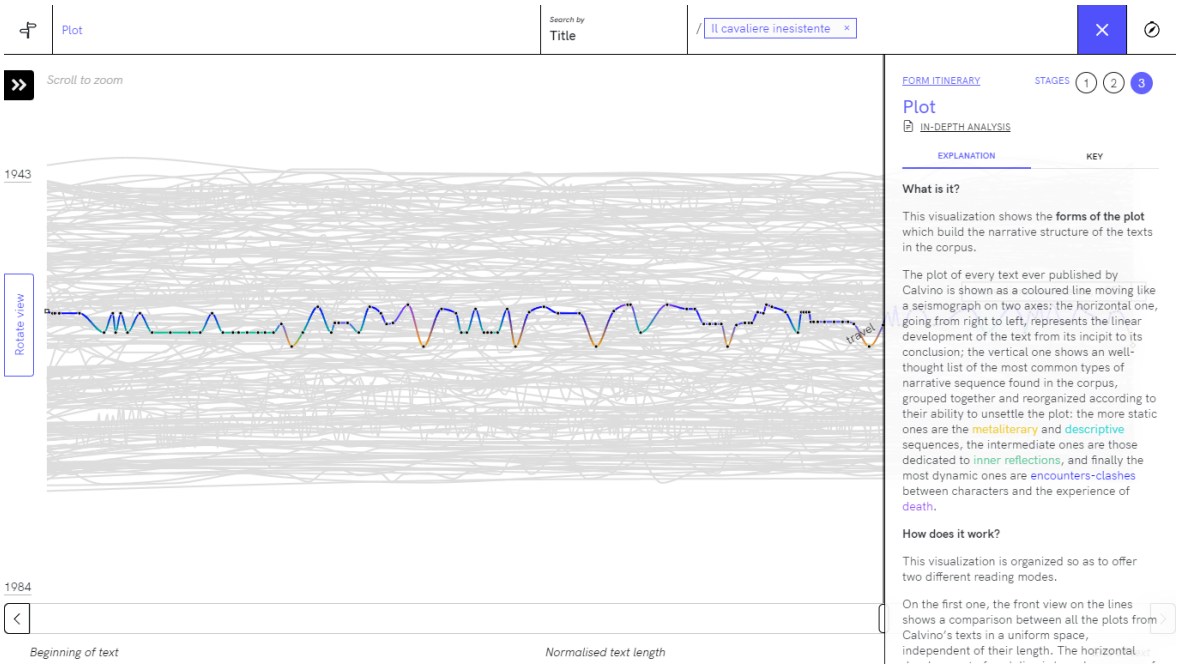

Fig. 3: Screenshot of the *Atlante Calvino* interface, which displays the shapes of the plots of his narrative works using line graphs

Network graphs can also be similarly used and serve as an excellent method for accessing information and are widely used in other fields. An example includes knowledge graphs and citation networks. Knowledge graphs are a type of database structured as a graph designed to gather and communicate knowledge (*Hogan et al., 2021, p. 3*). These databases serve as the foundation for search engines like Google and their algorithms for accessing and sorting web pages, such as PageRank (*Uyar & Aliyu, 2015, pp. 197-199*). Citation networks, on the other hand, are built through the interconnections among research within a given discipline, allowing researchers to identify influential articles, theories, and scholars, those most frequently cited, that have significantly contributed to the development of the field (*Gustafsson et al., 2014, p. 622*). This approach would facilitate meta-research, referring to the examination of research processes themselves, particularly through the reuse of character interaction networks as an alternative means of accessing the work. This can provide new insights into literary works while merging qualitative and quantitative methods. Researching a case study for the application of this would also contribute to applying Character Network Analysis to literary works with unconventional structures. Most projects to date have focused on plays, comedies, tragedies, and novels, traditional forms of literature with well-established characteristics and research traditions. Recent studies, such as those applying CNA to manga (e.g., *Sugishita & Masuda, 2023*), have shown promising results. For this reason, Character

Network Analysis will be applied to the *Spoon River Anthology* to examine its narrative structure through character interactions. Its composition of interconnected poems provides an initial perspective on what can be studied by reengaging with the text through the network of poems. The anthology will initially be analysed using Character Network Analysis to fill gaps in the existing literature on the subject, aligning with the interests of humanists who study both the author and the work. To evaluate the use of the character network as a tool for reaccessing the work, the anthology will be re-examined by revisiting the poems with recurring internal structures within the character network, analysing them through close reading as any humanist studying the work with this network would. It will be also re-analysed using an automatic analysis methodology, such as Sentiment Analysis, which has been recently adapted for narratology (*Ciotti, 2023, p. 105*). When combined with Network Analysis, this approach can reveal patterns of interaction linked to specific sentiment to highlight important sections of the literary work (*Min & Park, 2019, p. 6*). This will simulate the reuse of datasets constructed for network analysis by a humanist with digital skills who wishes to use the character interaction network for their specific research. Despite this interest, the research will follow a framework and structure typical of the hard sciences, with a focus on methodology, analysis, and discussion of the results. The work will then be presented to interested parties, engaging with scholars currently studying the author to gain not only the perspective of a Digital Humanist but also that of a potential humanist interested in using network visualisation for re-engagement with the literary work. Following this opportunity, a repository page will be created and published using GitHub Pages, allowing users to interact with the visualisation of the character interaction network. The goal is not only to provide a preliminary visualisation to firsthand verify its use but also to develop a repository where users can download the datasets and utilize them, promoting the dissemination of the work while ensuring accessibility even for those without digital skills, enabling them to conduct their own research. This will be achieved by offering tutorials within the repository for conducting initial analyses, accompanied by useful resources on the topic.

2. CHARACTER NETWORK ANALYSIS OF SPOON RIVER ANTHOLOGY

The *Spoon River Anthology* is a collection of poems by Edgar Lee Masters, with its final edition published by the Macmillan Company in November 1916 (*Hartley, 1963, pp. 3-4*). The work is notable for its editorial history, distinctive features, and its impact on contemporary literature, as well as its unique reception. The anthology's development began with support from Masters' friend William Marion Reedy, who first published the poems in his literary magazine, *Reedy's Mirror*, starting in May 1914. Masters used the pseudonym Webster Ford, and the poems continued to appear in the magazine until May 1915 (*Hallwas, 1992, pp. 2*-3). The collection gained such popularity that by mid-May, shortly after the final poem's appearance, the Macmillan Company had already prepared the bound edition for printing (*Narveson, 1962, p. 7*). The effort took a significant toll on Masters, who fell ill immediately after submitting the manuscript to Harriet Monroe for final editing (*Gregory & Zaturenska, 1946, p. 229*). The first edition was published in 1915, featuring 212 poems and concluding with *The Spooniad*. The 1916 edition included an additional 32 poems and an epilogue (*Hartley, 1963, pp. 4-5*). While the writing and publication of the anthology are relatively straightforward, the events that followed its release are more notable. The work has a strong presence in Italy, largely due to Fernanda Pivano's translation for Einaudi and the lasting influence of Fabrizio De André's concept album *Non al Denaro né all'Amore né al Cielo* (*VanWagenen, 2019, p. 680*). Despite the difficulties it faced in its home country, it remains a best-selling poetry collection in Italy. In contrast, Masters quickly faded from prominence in the United States, despite the initial success and impact of his work. He was seen as controversial and stylistically conservative, making him unable to sustain public interest and eventually harming his legal career (*Flanagan, 1974, pp. vii-viii*). This decline is still evident today, with Masters often marginalized within academic circles and relegated to regional literature. The difficulty of conveying the poem collection's textual fluidity online further complicates access and study (*Utard, 2024, pp. 20-21*). Even though it is claimed that Masters' work has not been fully understood by Italian readers (*VanWagenen, 2019*), the same sentiment also applies to American scholarship. These issues are reflected in the introductions of both the 1992 annotated edition edited by John Hallwas and the 1968 edition of *The New Spoon River* introduced by Willis Barnstone (*Hallwas, 1992 ;Barnstone, 1968*) Recently, Jason Stacy has provided a comprehensive examination of these issues and the real influence of

the poet, laying the groundwork for a renewed examination of his work and his figure (*Stacy, 2021*). It's more than appropriate to respond to the call for a rediscovery, drawing on the words of Martin Paul Eve, who in the conclusion of *The Digital Humanities and Literary Studies*, writes: "*Digital methods give us a way to reappraise literary works. These methods allow us to appreciate texts again, in the light of new knowledge*" (*Eve, 2022, p. 154*). Excluding *The Spooniad* and *The Epilogue*, which use different prosody, the anthology contains 244 free verse poems, beginning with *The Hill*, an incipit poem in the ubi sunt style and edited in a different font from the rest of the anthology (*Murphy, 1988, pp. 16-17*). The poems, inspired by epigrams from the *Greek Anthology*, suggested to him by William Marion Reedy, and by cemetery inscriptions from rural areas of the Midwest (*Gullo, 2019, pp. 407-408*), can be viewed as spiritual reflections or life summaries of the 244 characters from the fictional town of Spoon River. The town is named after the river in Illinois, although the characters and settings are inspired by real-life people and locations around the Spoon and Sangamon Rivers, particularly Petersburg and Lewistown. These settings reflect the 1880s, the period in which the anthology is set and during which the poet spent his adolescence in those places (*Chandler, 1921, pp. 252-254; VanWagenen, 2019, p. 684*). The titles of the poems typically represent the name of a deceased person, akin to a cemetery headstone, with the text presenting the character's final lament after death. Since all characters are from the same fictional town, many discuss events that link various poems, each from their own perspective, which makes the narration of events subjective. The events in the Anthology are not presented in chronological order. For instance, poem number 51, *Lois Spears*, refers to *Willard Fluke*, the father of the narrator, whose poem is positioned as number 53. Despite this, the events in *Willard Fluke* occur before those in *Lois Spears* due to the father's death preceding the daughter's. This interconnectedness, or what May Swenson termed the "*lurid web of small-town scandal*" (*Swenson, 1962, p. 12*), highlights a shift in American literature. Previous literature idealized rural life, whereas this work presents a more realistic view. This theme, sometimes called the "Revolt from the Village", contrasts romanticized with a realistic portrayal of small-town life (*Marcia, 2021, pp. 7-8*). This realism creates a geographically isolated area marked by pre-industrial stagnation, contrasting with an idealized vision of the rural past (*Carroll, 2021, pp. 118-121*). Masters deconstructs the myth of the small town as an ideal community of innocence, simplicity, and trust, presenting a more complex reality (*Chandran, 1988, p. 439*). He constructs a detailed microcosm with nearly every

occupation known in the Midwest, using real names that are either shuffled or slightly altered (*Flanagan, 1953, p. 234*). His inspiration from real life is evident, as the anthology references numerous actual events (*Burgess, 1987, p. 348*). Building on the earlier discussions and research questions, those interactions between poems/characters need to be studied by applying Character Network Analysis to examine the character network in Spoon River Anthology, given its vastness. This approach provides an overview of the interpoetic narrative structure, shaped by the characters' first-person accounts, while testing this methodology on a literary work that deviates from traditional forms. The study also includes an analysis of changes in the interpoetic narrative structure between the first and second editions of the anthology to trace what the new poems contributed to the interpoetic narrative. To further explore the potential of network analysis for re-engaging with the text through both quantitative and qualitative methods, a topological analysis of the internal network structures was conducted. Following this, the poems were examined comprehensively, guided by insights from the network topology, which will be detailed in the results section. A mixed-methods approach was then employed, using Sentiment Analysis to explore the sentiment directed toward characters inspired by Lewistown residents. The research questions guiding the validation of my initial hypotheses on CNA are:

- How is the interpoetic narrative within the anthology structured and organized, especially regarding the interactions between poems, the topology of these interactions, and their placement in the anthology?
- What changes did the 1916 edition introduce to the interpoetic narrative, specifically in terms of poem interactions and their placement?
- How do characters inspired by real Lewistown residents interact with other poems and characters, and how is the topology of these interactions structured?
- How are the sentiments of other characters interacting with those inspired by Lewistown residents portrayed, and how is sentiment, particularly negative sentiment, distributed among these characters and positioned within the anthology?

The analysis is divided into three sections: a methodological section explaining how data were extracted, modeled, and analysed, alongside the strengths and limitations of the approach; a results section presenting and analysing the findings; and a discussion section

that will contain interpretations of the results and insights for the reassessment of the work thanks to them.

2.1 Methodology

The methods section is organized into distinct phases: data extraction, collection and modeling, the metrics, visualisations, and tools used for analysis, as well as the limitations and strengths of the employed methodology.

2.1.1 Data extraction, collection and modeling

The analysis required data on the poems/characters and their interactions, their presence across different editions, and whether they were inspired by Lewistown residents. To streamline the data and metadata collection, two databases were initially created: one compiling the poems and their attributes, and another documenting the interactions and their related details. The data were manually drawn from two sources: the 1992 annotated edition by Hallwas, which was used for analysing the Anthology, including interactions, poem placements, and character inspiration from real individuals; and a version of *Spoon River Anthology* that categorizes the poems by edition, accessed via the Wayback Machine, a digital archive provided by the Internet Archive. The extraction and collection of character interactions followed the framework established by Labatut and Bost in their 2019 survey on Character Network Analysis but adapted to the Anthology structure (*Labatut & Bost, 2019*). Close reading was necessary for analysing mentions, given the lack of prior studies in this area and limitations of CNA exposed in the first Chapter. There was no existing database or established understanding of how the connection of characters within the work occurred. To ensure the integrity of the research, all poems were analysed except for *The Spooniad* and *The Epilogue*, as these possess distinct characteristics that differ significantly from the other poems. The analysis identified that interpoetic narration primarily happens through character mentions in the poems. Three types of mentions were classified: "Direct," where another character is cited by the title of their poem; "Semi-Direct," where only part of the poem's title is referenced; and "Indirect," where no elements of the title are used, but the referenced poem can be inferred through contextual knowledge.

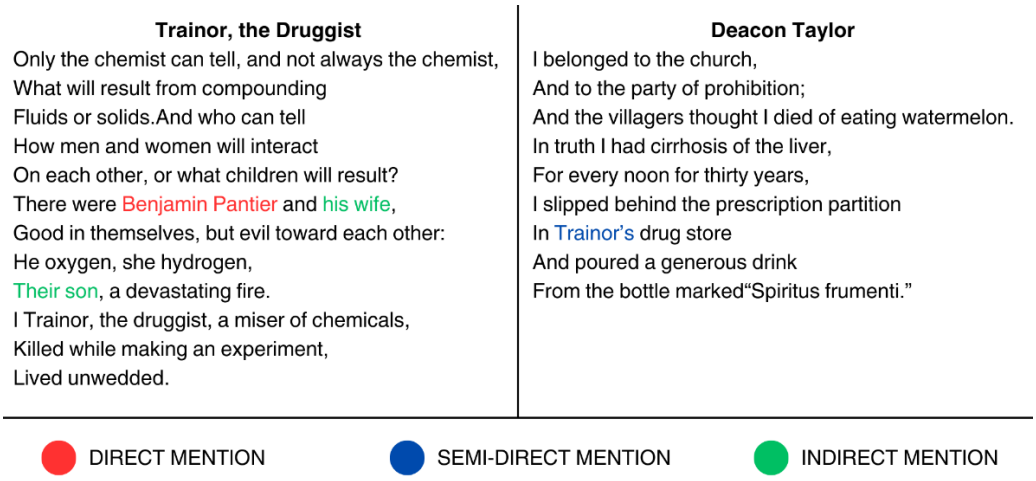

Fig. 4: Example of a direct, semi-direct, and indirect reference based on the poems *Trainor, the Druggist* and *Deacon Taylor*

Indirect mentions that could not be conclusively linked to a specific poem were excluded from the analysis. When a poem or character was referenced through another character, it was considered a double mention (e.g., in the poem *'Butch' Weldy*, where "Old Rhodes' son" is mentioned). Since no previous study had thoroughly examined these types of mentions, and their variations could complicate automated extraction, the decision to manually enter the data into the databases was ultimately maintained. The same process was applied to the extraction of auxiliary information from all sources, such as the creation of an ID for each poem to study its position within the work. To extract and analyse sentiment towards characters inspired by Lewistown residents, interactions featuring these characters were first retrieved from the interaction preliminary database. Sentiment analysis was then conducted semi-automatically using ChatGPT-4o mini, guided by a prompt (*Appendix B-1*) designed to fit the tool's capabilities and limitations (*Kheiri & Karimi, 2023*), with acknowledgment of its occasional errors in sentiment classification (*Kamil & Shah, 2024*). The sentiment data were compiled into a third database, categorizing interactions and sentiment values as Negative, Neutral, or Positive. This database recorded the final sentiment values and compared them with those suggested by ChatGPT-4o mini. It also included columns explaining the rationale for each analysis and notes on whether the analysis was accurate or needed re-evaluation and correction, with explanations provided. The workflow is illustrated in Figure 5.

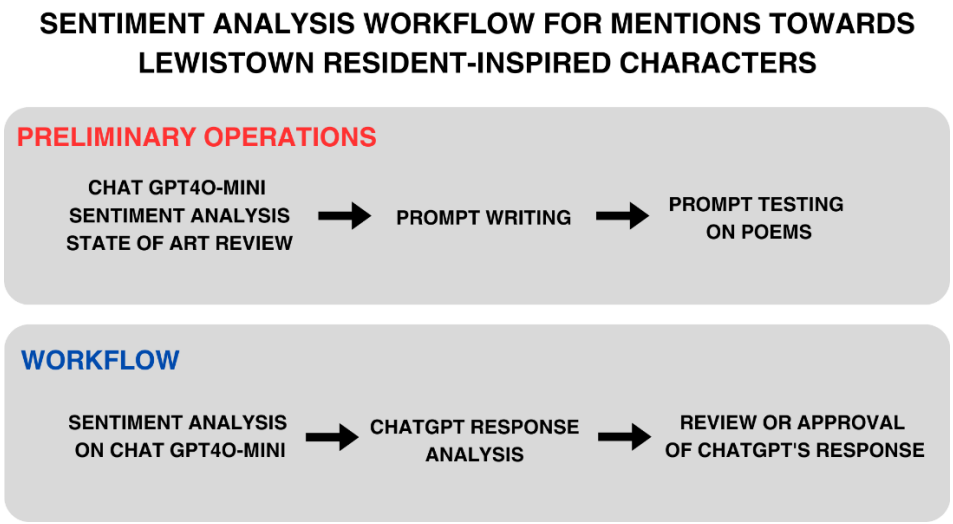

Fig. 5: Illustration of sentiment Analysis workflow for mentions towards Lewistown resident-inspired characters

The data from the poem and interaction databases were organized to facilitate analysis. In the poem database, poems column were modeled as nodes, with columns detailing their attributes. Interactions were represented with a Source and Target designation: the Source node was the poem making the mention, and the Target node was the poem being mentioned. Each interaction included relevant attributes for the study. To prepare the data for analysis in Python, databases were imported into a directed network using the Pandas and NetworkX libraries, within a Jupyter Notebook developed in Visual Studio Code. A directed network was used because the mentions represent a one-way interaction between poems, rather than a mutual relationship (*Labatut & Bost, 2019, pp. 19-20*). This setup allows nodes to represent both poems and characters simultaneously. The complete workflow is shown in Figure 6. All datasets are available in GitHub repository found in *Appendix A*

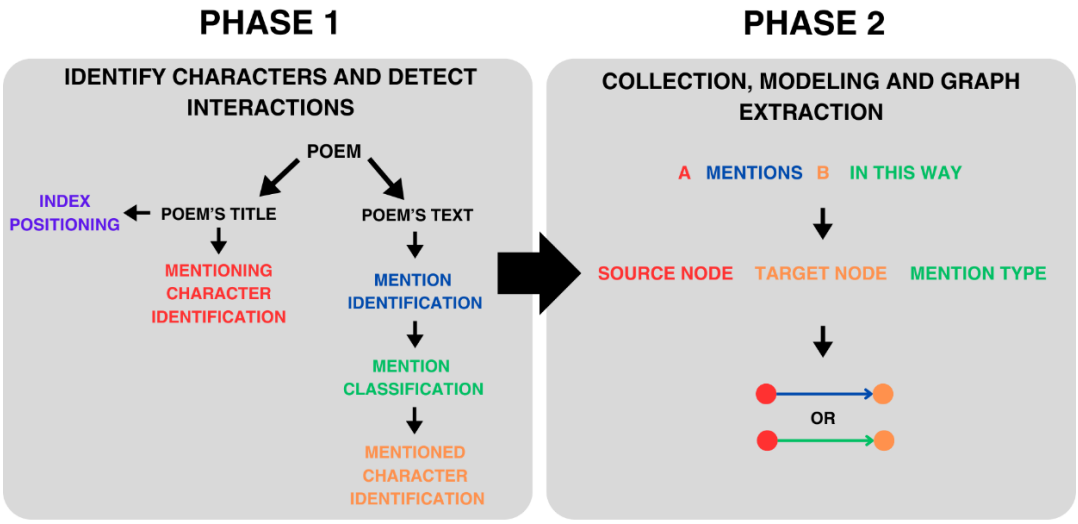

Fig. 6: Illustration of mention extraction workflow for *Spoon River Anthology* created by adapting the generic character extraction process by Labatut and Bost to the work

2.1.2 Metrics, visualisations, and tools used for analysis

To investigate the narrative structure of the *Spoon River Anthology* through Character Network Analysis, this study utilized a combination of network visualisation tools and key network metrics to handle the complexity of the interpoetic narrative. Network visualisations were generated using the *PyVis* library, which supports real-time adjustments to the network's physics, ensuring the visualisation met the necessary characteristics and integrated graphs created with *NetworkX*. Customization of edge and node colors was done based on the analysis requirements and selected attributes. Metrics related to node and edge counts, in-degree centrality, out-degree centrality, Betweenness Centrality, PageRank, density, and average clustering coefficient were calculated using functions provided by *NetworkX*. Classic metrics, including node and edge counts and in-degree and out-degree, offered insights into the prominence of characters within the network. They highlighted whether a character was frequently mentioned by others or was influential in invoking the presence of other characters. Density and average clustering coefficient provided an overview of the network's overall connectivity and assessed potential clustering among characters based on the frequency and distribution of mentions. Along with standard network metrics, Betweenness and PageRank centralities were employed to evaluate the bridging roles of characters across multiple

stories and to determine if frequently discussed characters were well-connected. Distribution graphs for all four centrality measures were created using the *NumPy* and *Matplotlib.pyplot* libraries, which are compatible with *NetworkX*. To analyse the placement of poems and their interactions within the anthology, including their attributes, *NumPy* was used to create the adjacency matrix, and *Matplotlib.pyplot* was employed for visualisation and legends. The matrix axes were aligned with the poem Ids to preserve the poems' order as listed in the index. Supplementary data on nodes, edges, and attributes were printed, tabulated, or calculated using straightforward functions, such as those for averaging. For studying internal structures, specifically topological analysis, simple structures were described, and triads were classified using the triads_type function according to Snijders' 2012 classification of triads (*Snijders, 2012*). Internal structures were identified by analysing initial metrics, and a specific code was developed to detect structures with above-average in-degree and out-degree values and hierarchical connections, as suggested by preliminary results from the adjacency matrix (*Appendix B-2*). These structures were extracted, classified, and examined through visualisations generated with PyVis. Final visualisations were refined using Canva's image editor, and their data were tabulated based on the visual outputs.

2.2 Results

The results section is organized according to the research questions addressed, with each part following the research workflow. In this section, the term "Mention-Network" refers to the network of interactions involving mentions between characters.

2.2.1 Analysis for Interpoetic Narrative

The network visualisation of the mention interactions in Figure 7 revealed 18 distinct connected components. These included one large component with the majority of nodes and edges, primarily located in the center, and 17 smaller components. Metrics from the Mention-Network are detailed in Table 1, and metrics from the largest component are outlined in Table 2.

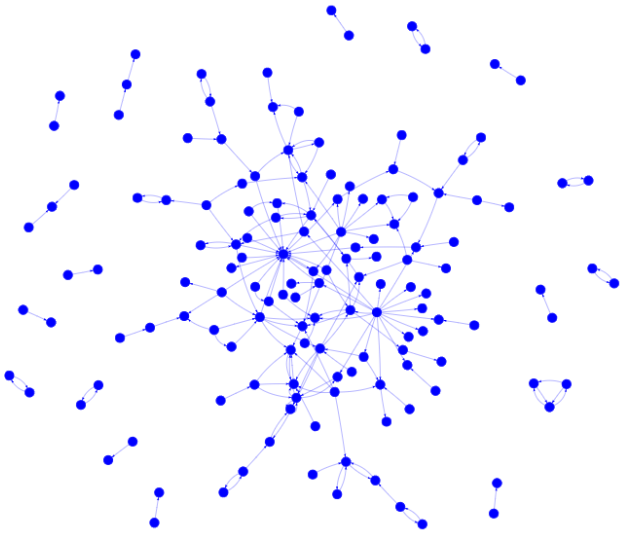

Fig. 7: Visualisation of the Mention-Network in the Spoon River Anthology

Metrics of the Mention-Network	**Value**
Nodes count	144
Edges count	196
Connected components	18

Tab. 1: Metrics extracted from the Mention-Network

Metrics of the Largest connected component	**Value**
Nodes count	107
Edges count	168
Density	0.0148
Average clustering coefficient	0.0734

Tab. 2: Metrics extracted from the largest connected component

The histogram for indegree distribution in the largest connected component, shown in Figure 8-A, indicates indegree values ranging from 0 to 17. There is a notable frequency gap for indegree values between 7 and 16. The highest frequency of 40 occurs at an indegree value of 1, with another peak at an indegree value of 0, slightly above 30. The distribution shows a strong right skew, with a frequency drop of 25 between indegree values of 1 and 2, continuing until it reaches the highest indegree value. Figure 8-B displays the outdegree distribution histogram, with outdegree values from 0 to 17. Frequency gaps are observed between values 6 and 9, and between values 11 and 16. The highest frequency, just over 45, is at an outdegree value of 1, which is more pronounced compared to values 0 and 2, each having a frequency around 20. The outdegree

distribution is also right-skewed, with an asymmetric drop at an outdegree value of 4. The corresponding indegree and outdegree data are detailed in Table 3.

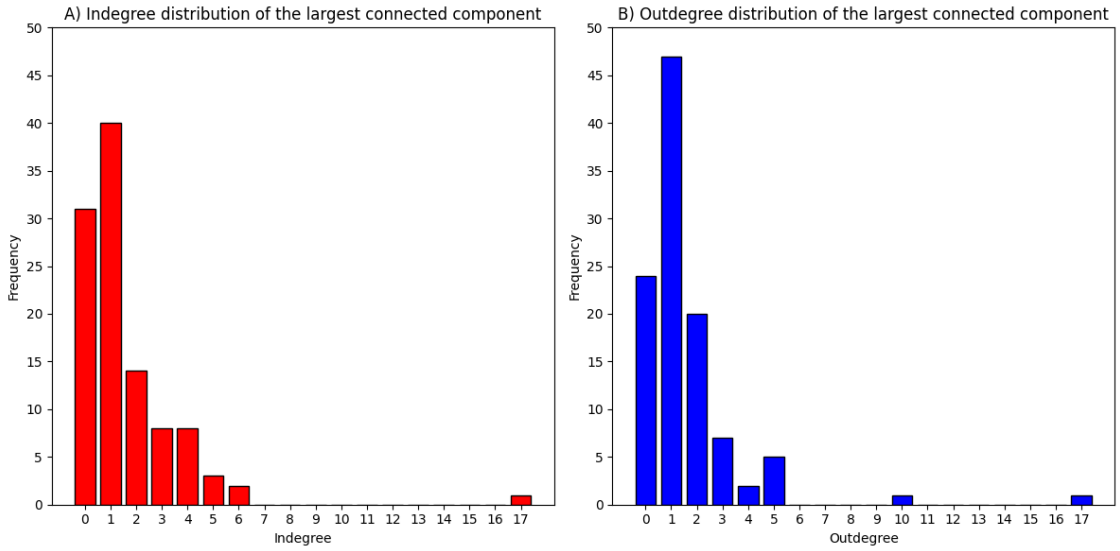

Fig. 8: Histograms representing (A) the indegree distribution and (B) the outdegree distribution by node frequency and indegree/outdegree values for the nodes of the largest connected components in the Mention-Network

Metrics regarding indegree and outdegree of the largest connected component	**Value**
Average indegree and outdegree of nodes	1.570
Number of nodes with indegree above average	36
Number of nodes with outdegree above average	36
Maximum indegree	17
Maximum outdegree	17

Tab.3: Summary of metrics extracted from the indegree and outdegree of nodes in the largest connected component of the Mention-Network

The histogram of PageRank values for nodes in the largest connected component, shown in Figure 9-A, ranges from 0.00 to 0.08 in 0.01 intervals. Noticeable gaps appear between 0.03 to 0.04 and 0.06 to 0.07. The highest frequency, just over 80, is found in the 0.00 to 0.01 range, with a second peak between 0.01 and 0.02, showing a drop of around 65 in frequency compared to the first range. The distribution is right-skewed, with a slight increase in frequency observed between 0.07 and 0.08. Figure 9-B presents the histogram for betweenness centrality values, ranging from 0.00 to 0.04 in 0.01 intervals. The peak frequency, slightly under 95, is in the 0.00 to 0.01 range, followed by a sharp drop in the

next interval, with a frequency around 6, and even lower values exhibit a right-skewed distribution. Summary data for PageRank and betweenness centrality are provided in Table 4.

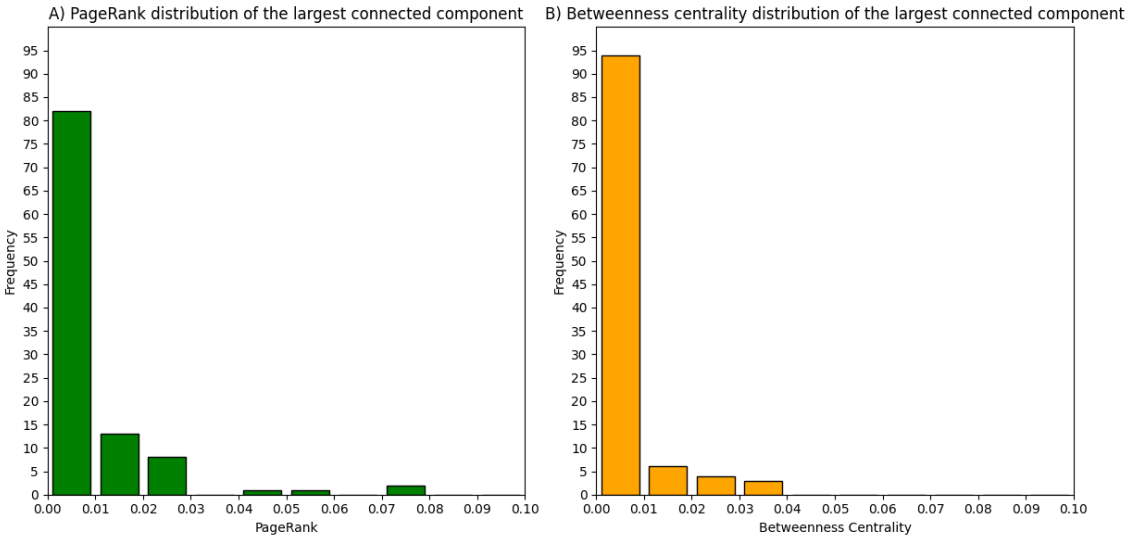

Fig. 9: Histograms representing (A) the PageRank distribution and (B) the Betweenness Centrality distribution by node frequency and value for the largest connected components in the Mention-Network

Metrics regarding PageRank and Betweenness centrality of the largest connected component	**Value**
Average PageRank	0.009
Average Betweenness centrality	0.003
Number of nodes with PageRank value above the average	25
Number of nodes with Betweenness centrality values above average	13
Maximum PageRank	0.076
Maximum Betweenness centrality	0.039

Tab. 4: Summary of metrics extracted from the PageRank and Betweenness centrality values of nodes in the largest connected component of the Mention-Network

All data from the mention type attribute of edges are presented in Table 5.

Mention type attributes data	Value
Number of Direct Mentions	118
Number of Semi-Direct Mentions	52
Number of Indirect Mentions	26
Number of nodes with only Direct incoming edges	66
Number of nodes with only Semi-Direct incoming edges	37
Number of nodes with only Indirect incoming edges	20
Number of nodes with all three types of incoming edges	5

Tab.5: Summary of mention type attribute data

The visualisation of the adjacency matrix for the Mention-Network, organized by the poem's appearance Id in the Anthology and with dots colored by mention type, is shown in Figure 10. Most points are concentrated within the Source Poem range of 0-160 and the Target Poem range of 0-140, where the majority of Semi-Direct mentions are also located. The matrix reveals various patterns and clusters. A pattern of nodes along the primary diagonal starts to dissipate around Target Poem 140. Four ranges show no mention interactions: around 30-40 and 40-50, with only two interactions separating them, and in the ranges of 80-90 and 140-160. Three clusters are observed near the primary diagonal pattern, specifically between Poems 10-30, 50-70, and 90-110. Most mentions are below the primary diagonal, with the majority of indirect mentions clustered around this diagonal pattern. A corner pattern appears with an angle in the cluster of nodes around the primary diagonal in the range 10-30. A vertical pattern extends along the Source Poems axis, ending just before poem 230, while a horizontal pattern follows the Target Poems axis, extending slightly beyond poem 140. Another vertical pattern targets poems in the range 50-80 and extends more diffusely across the Source Poems axis until just before poem 230. A prominent straight line of points runs along the Source Poems axis in the range 100-110 of the Target Poems, dividing the matrix into two parts, with most interactions on the left side. Data on the distance between characters in the literary work, according to their mention type, are compiled in Table 6.

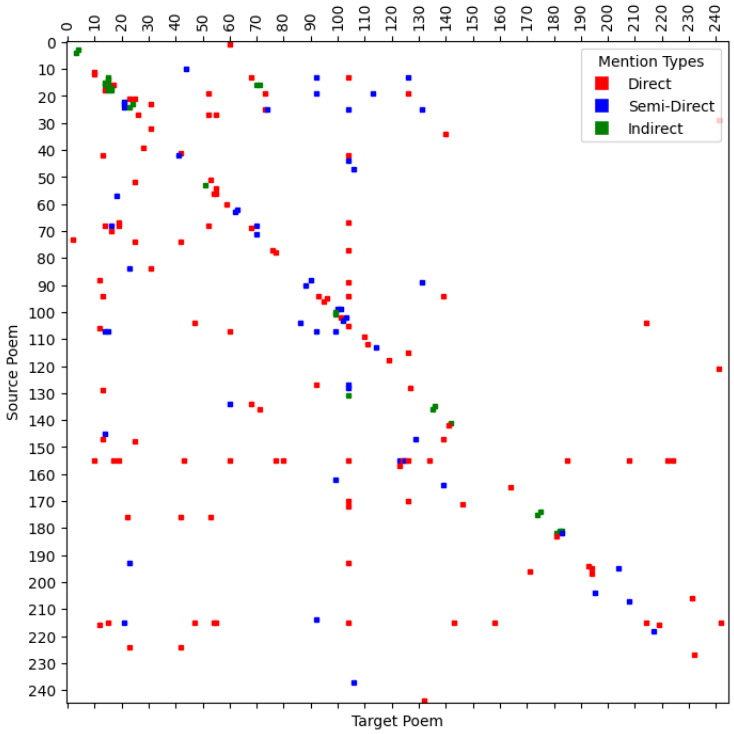

Fig. 10: Visualisation of the adjacency matrix of the Mention Network with axes order based on the position of poems in the Anthology and points colored according to the type of mention

Mention Type	**Average distance between poems**
Direct	53.25
Semi-Direct	43.19
Indirect	6.42

Tab. 6: Average distance between poems based on Mention-Type attribute

The smaller connected components in Figure 1 displayed varied structures, which have been quantified and classified in Table 7 and illustrated in Figure 11.

Smaller connected component structures	**Quantity**
Dyad (one edges)	9
Dyad (Reciprocal edges)	5
Triad (Type 210)	1
Triad (Type 021U)	1
Triad (Type 021C)	1

Tab. 7: Classification and quantity of smaller connected component structures of the Mention-Network

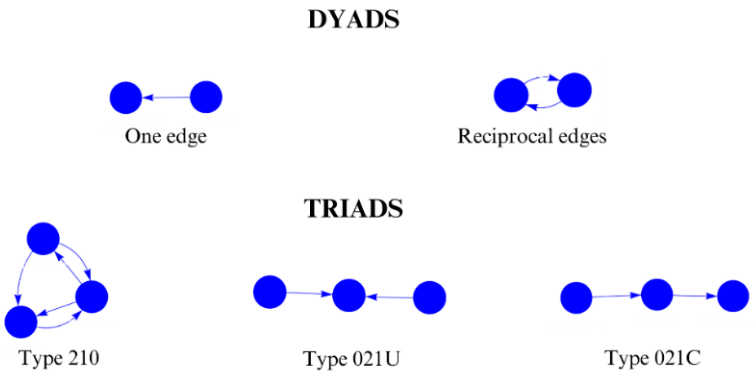

Fig. 11: Visualisation of classified smaller connected component structures found in the Mention Network

The data collected from the network and the adjacency matrix visualisation indicated connected hierarchical structures. The analysis of the largest connected component identified "Hub-and-Spoke" structures[1]. These structures, along with their metrics, are detailed in Table 8 and illustrated in Figure 12, which shows the role of each node within the structure, categorized into Hubs, spokes shared by at least two structures, and unique spokes.

Metrics of the "Hub and Spokes" structures	**Value**
Number of Hubs	11
Number of shared spokes between structures	15
Number of not shared spokes	31
Average nodes in the structures	9
Average edges in the structures	12
Number of nodes in the smallest structure	5
Number of edges in the smallest structure	5
Number of nodes in the biggest structure	21
Number of edges in the biggest structure	28
Number of nodes within a "Hub and spokes" structure	57

Tab. 8: Summary of metrics from the "Hub and spokes" structures found inside the largest connected component of the Mention-Network

[1] The Hub-and-Spoke model is commonly applied in industrial contexts, particularly in the development of transportation and telecommunications systems. In this network, all points are interconnected, with key locations, referred to as hubs, serving as central routing points for information or materials. By channeling flows from smaller nodes, or spokes, through these hubs, the system increases efficiency, making better use of resources such as equipment and labor (*An et al., 2015, p. 103*).

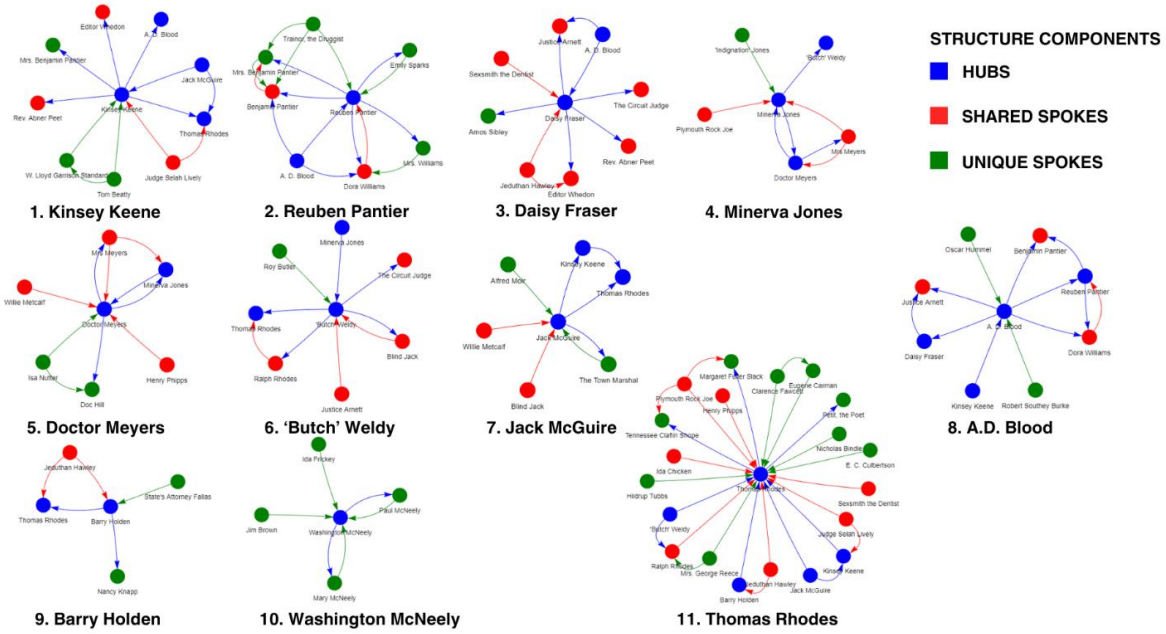

Fig. 12: Visualisation of the Hub-and-spoke structures within the largest connected component of the Mention-Network ordered by hub position in the literary work and colored by function of the components in the structure

The visualisation of connections between hubs (Fig. 13) illustrates their formation into a connected component, excluding the hub for *Washington McNeely*. A central triadic structure (Type 030T) involving *Kinsey Keene*, *Jack McGuire*, and *Thomas Rhodes* is visible.

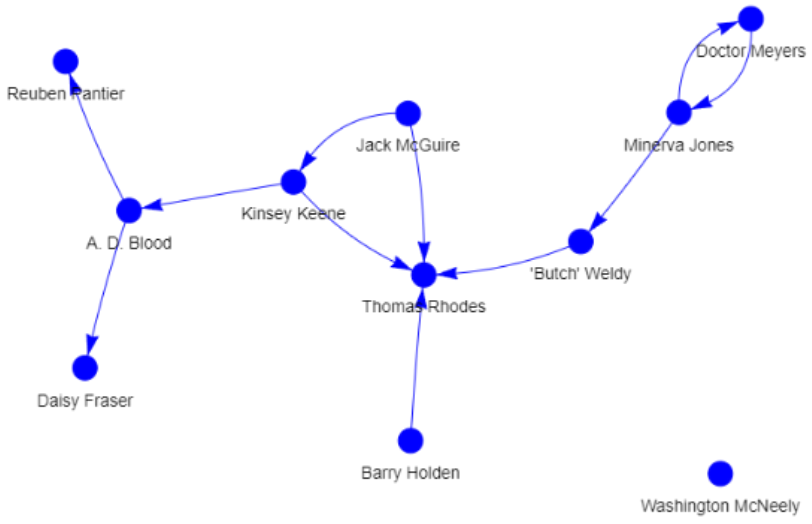

Fig. 13: Visualisation of the connections between hub nodes of the hub-and-spoke structures identified in the largest connected component of the Mention-Network

The connections between hubs and shared spokes create several triangular triads (Fig. 14), which are extracted and classified in Figure 15, with the data presented in Table 9.

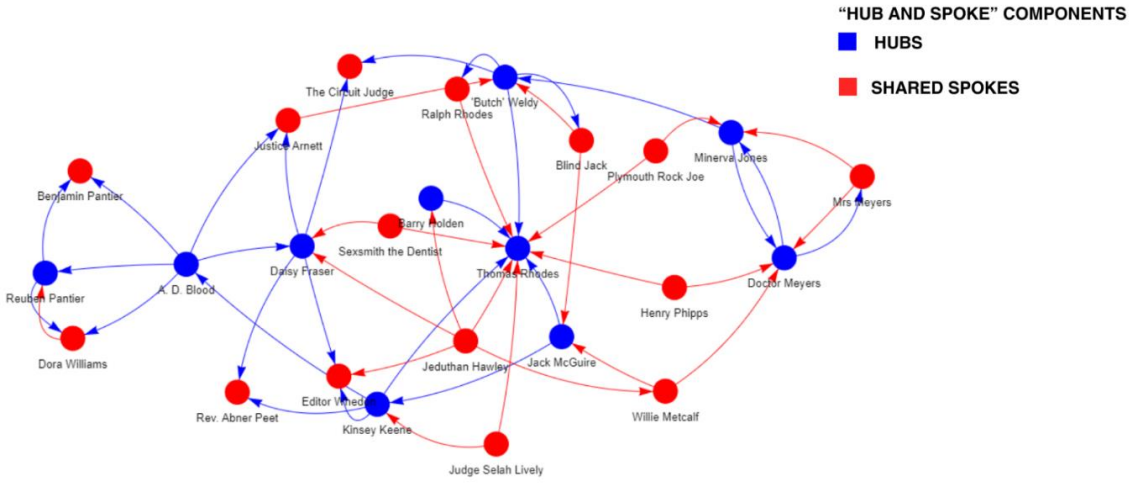

Fig. 14: Visualisation of the connections between hub nodes and shared spokes in the hub-and-spoke structures identified in the largest connected component in the Mention-Network colored by function of the node in the structure

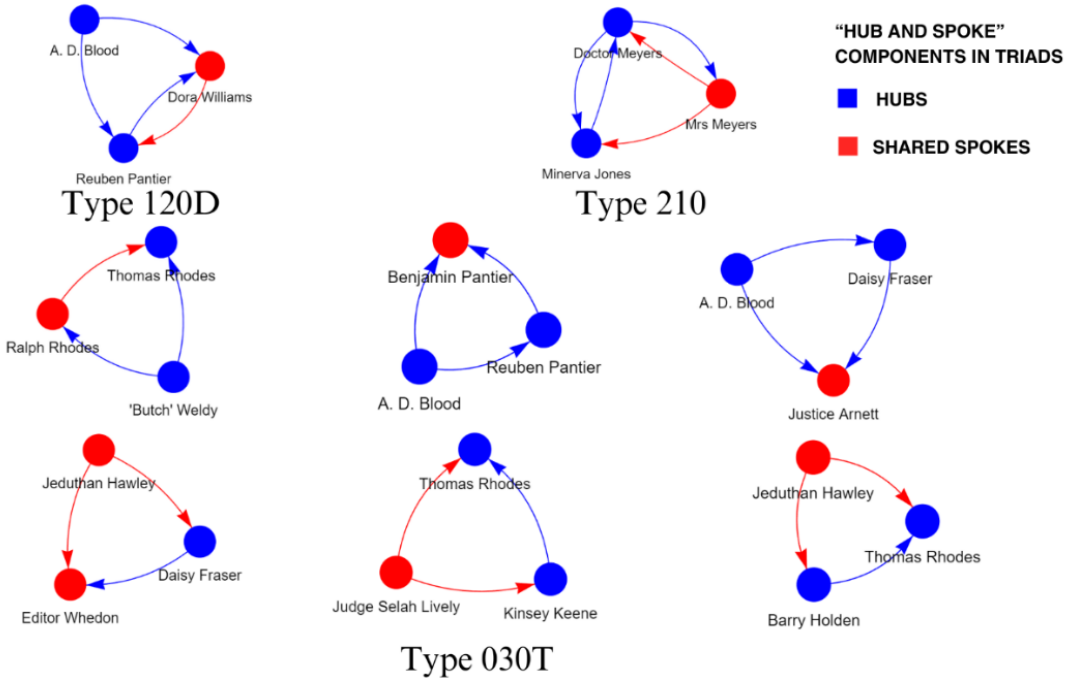

Fig. 15: Visualisation of the triadic structures formed by hubs and shared spokes in the hub-and-spoke structures of the largest connected component in the Mention-Network colored by function of the node in the structure

Data gathered from the triangular triads involving Hubs and Shared Spokes	Value
Number of type 030T triads	6
Number of type 120D triads	1
Number of type 210 triads	1
Number of triads with more than one Hub	7
Number of triads with more than one Shared Spoke	1
Number of Hubs in multiple triangular triads	3
Number of Shared Spokes in multiple triangular triads	1
Maximum number of triads to which the same Hub belongs	3

Tab. 9: Data extracted from the triangular triads involving Hubs and Shared Spokes and their value

The connection between all components of the "Hub and Spokes" structures (Fig. 16) demonstrates how *Washington McNeely*'s hub reconnects with other structures and forms new triangular triads. These new triads are extracted and classified in Figure 17, with their data presented in Table 10.

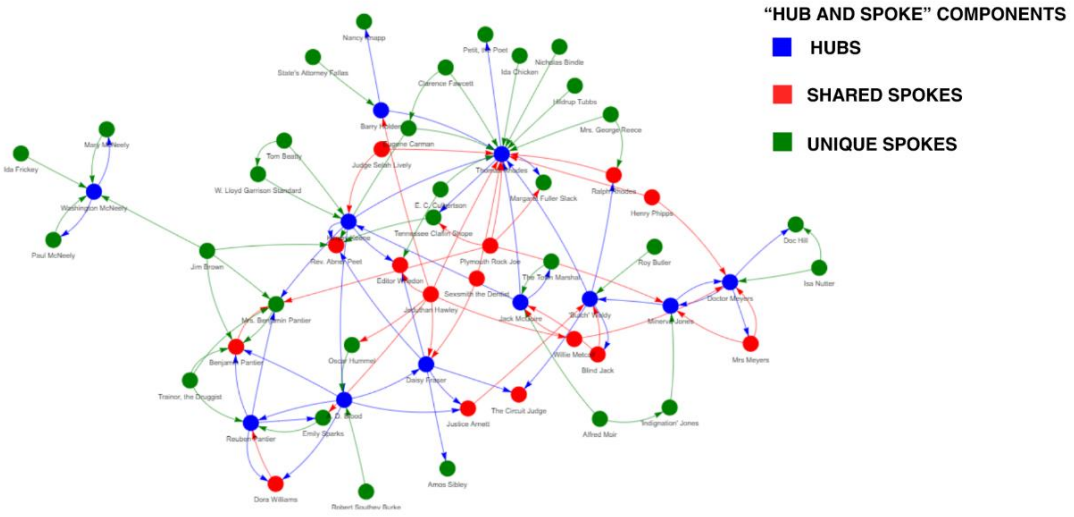

Fig. 16: Visualisation of the connections between hub nodes (in blue), shared spokes (in red) and non-shared spokes (in green) in the hub-and-spoke structures identified in the largest connected component of the Mention-Network

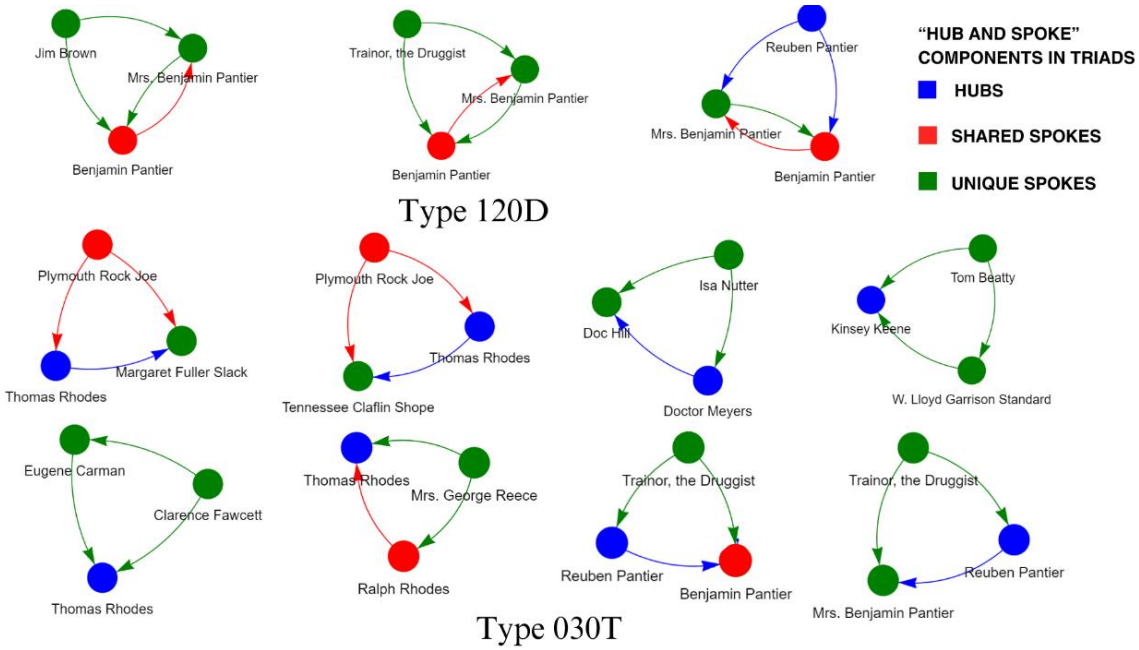

Fig. 17: Visualisation of triadic structures formed by hubs, shared, and non-shared Spokes in the hub-and-spoke structures of the largest connected component in the Mention-Network

Data gathered from the triangular triads between Hubs and Shared Spokes and Unique Spokes	**Value**
Number of type 030T triads	8
Number of type 120D triads	3
Number of triads with more than one Unique Spoke	7
Number of Hubs in multiple triangular triads	2
Number of Shared Spokes in multiple triangular triads	2
Number of Unique Spokes in multiple triangular triads	2
Maximum number of triads to which the same Hub belongs	3
Number of triads formed by all three components of the Hub and Spokes structures	5

Tab. 10: Data extracted from the triangular triads involving Hubs, Shared Spokes and Unique Spokes and their value

All the "Hub and Spokes" structures are centrally positioned in the Mention-Network, as shown in Figure 18, while most of the remaining nodes are located in peripheral areas.

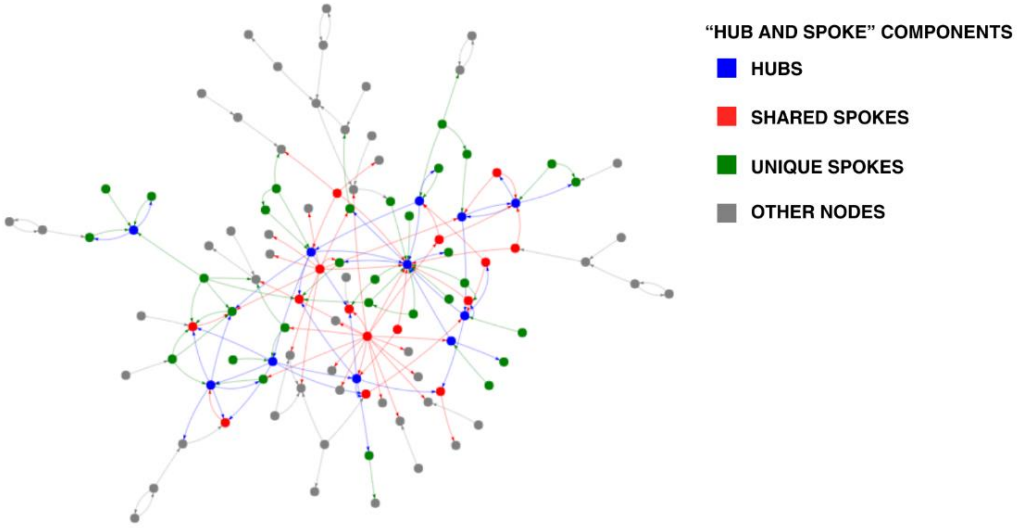

Fig. 18: Visualisation of the position in the Mention-Network of the hub-and-spoke structures showing hub nodes (in blue), shared spokes (in red), non-shared spokes (in green), and other nodes (in gray)

2.2.2 Change in interpoetic narrative with the 1916 edition

The network visualisation, with nodes colored by their edition presence (Figure 19), revealed that most nodes were within the largest connected component of the Mention-Network, while one node was in a dyad with a single edge. Only two poems from the 1916 edition have connecting edges. Metrics for the nodes of poems from the 1916 edition are summarized in Table 11.

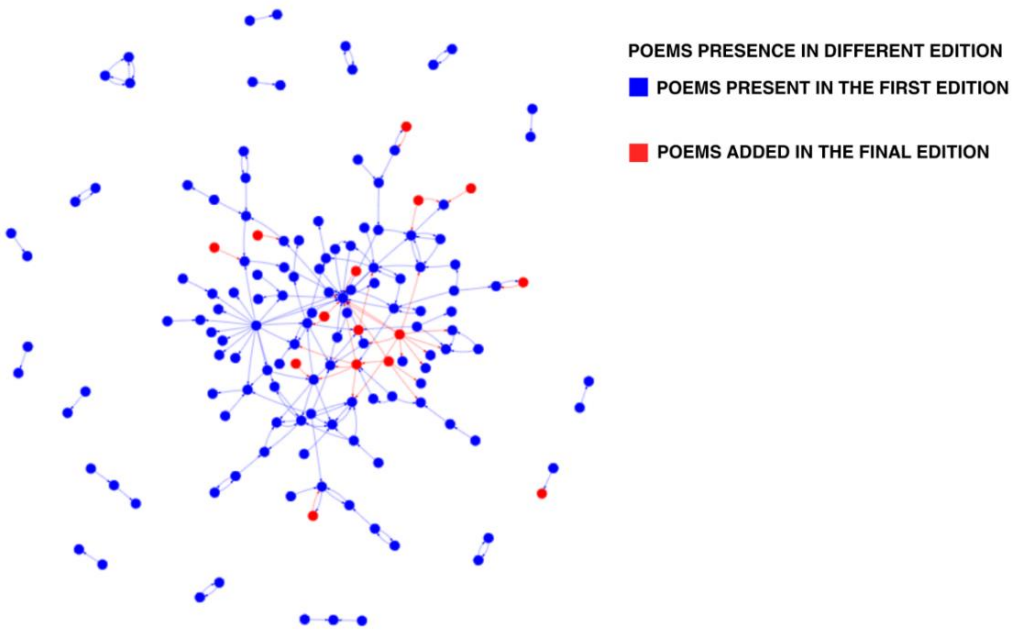

Fig. 19: Visualisation of the Mention-Network in the Spoon River Anthology with poems colored by edition appearance

Metrics of poems added in the 1916 edition	**Value**
Number of poems added in the final edition	33
Number of nodes in the Mention-Network	15
Number of nodes in the largest connected components	14

Tab. 11: Summary of metrics gathered from the nodes added in the 1916 edition and their presence in the Mention-Network

The histogram showing the distribution of indegree for poems in the final edition and within the largest connected component (Figure 20-A) indicates indegree values from 0 to 4, with the highest frequency of 9 at indegree 0. There is a gap at indegree 2, and higher indegree values each have a frequency of 1. Similarly, the histogram for outdegree distribution for these poems (Figure 20-B) shows the highest frequency of 9 at an outdegree value of 2. Gaps are noted at outdegree 3 and within the range of 6 to 9, while the highest outdegree values of 4, 5, and 10 each have a frequency of 1. Metrics for indegree and outdegree values for this set of nodes are presented in Table 12.

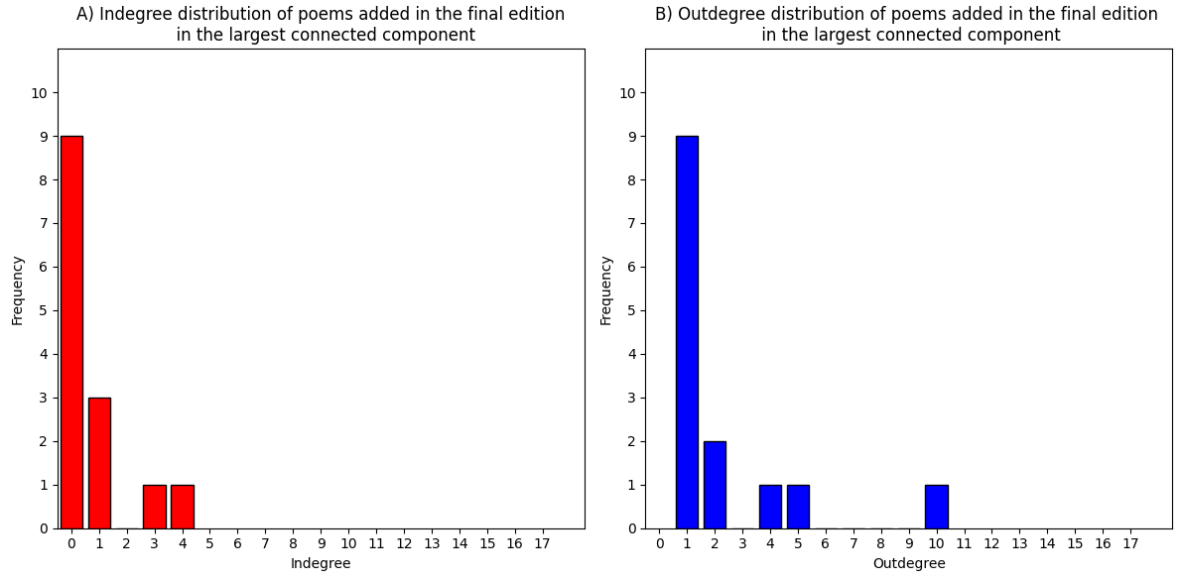

Fig. 20: Histograms representing (A) the indegree distribution and (B) the outdegree distribution of the poems added in the final edition belonging to the largest connected component

Metrics regarding indegree and outdegree of the nodes from poems added in 1916 edition	**Value**
Number of nodes of the largest connected component with indegree above the average	2
Maximum indegree for the largest connected component	4
Number of nodes of the largest connected component with outdegree above the average	5
Maximum outdegree in the largest connected component	10

Tab. 12: Summary of values extracted from the indegree and outdegree of nodes of poems added in 1916 edition

The histogram showing the PageRank distribution for the same sample of poems (Figure 21-A) displays values from 0.00 to 0.03, with the highest frequency in the 0.00-0.01 range. The next range has a frequency of 2, while the 0.02-0.03 range has a frequency of 1. The histogram for betweenness centrality (Figure 21-B) shows values from 0.00 to 0.04, with the highest frequency of 12 in the lowest range. Remaining values each have a frequency of 1 and fall within the 0.01-0.02 and 0.03-0.04 ranges, with a gap between these ranges. Metrics for PageRank and betweenness centrality for this set of nodes are listed in Table 13.

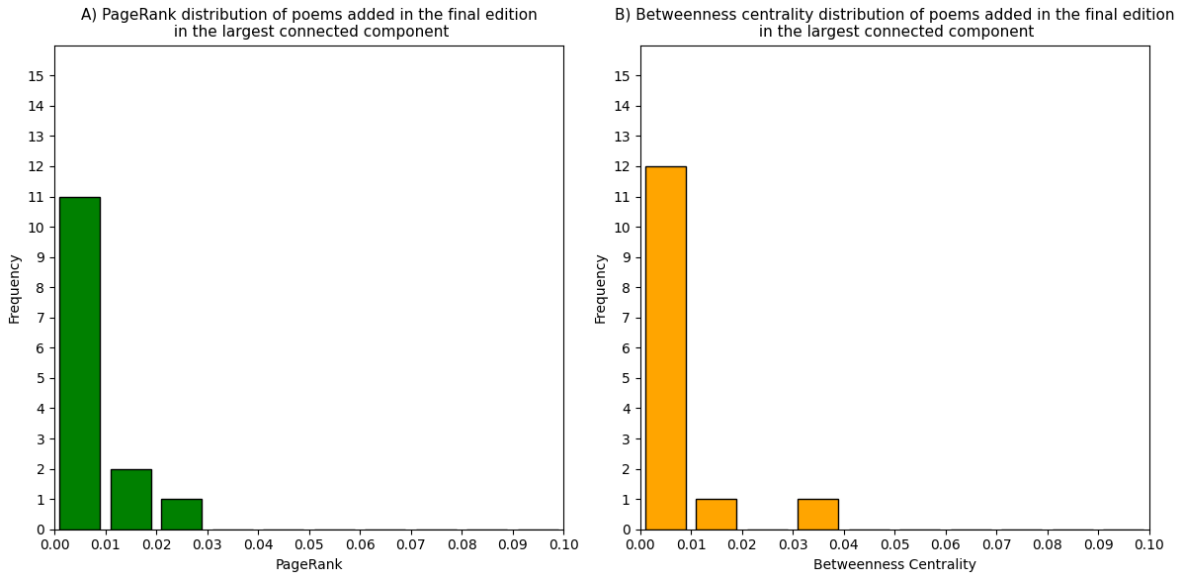

Fig. 21: Histograms representing (A) the PageRank distribution and (B) the Betweenness Centrality distribution of poems added in the final edition belonging to the largest connected component

Metrics regarding PageRank and Betweenness centrality values of the nodes from poems added in 1916 edition	**Value**
Number of nodes of the largest connected component with PageRank values above the average	3
Maximum PageRank in the largest connected component	0.021
Number of nodes of the largest connected component with Betweenness centrality values above the average	2
Maximum Betweenness Centrality in the largest connected component	0.030

Tab. 13: Summary of values extracted from the PageRank and Betweenness centrality of nodes of poems added in 1916 edition

The adjacency matrix of the Mention-Network, shown in Figure 22, highlights interactions among poems added in the final edition. Points are clustered around the diagonal line, and the matrix displays a corner pattern with source and target poems in the 10-30 range. The vertical axis represents Source Poems and ends just before poem 210, while the horizontal axis covers Target Poems up to just before poem 130. Two distinct straight lines are visible: one vertical between Target Poems 100-110 and one horizontal between Source Poems 210-220.

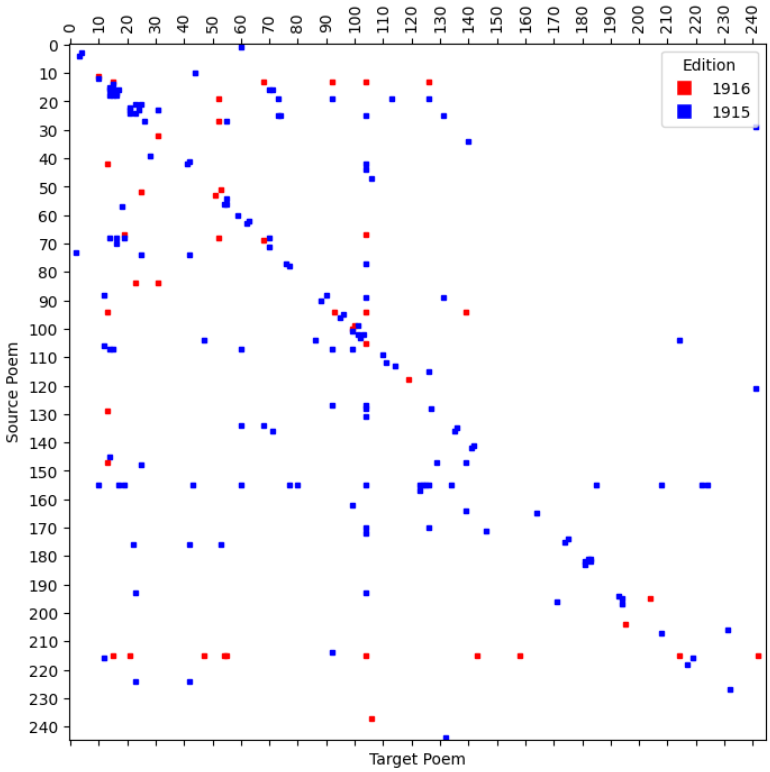

Fig. 22: Adjacency matrix of the Mention Network of the Spoon River Anthology, with axes ordered based on the appearance of the poem in and interaction colored according to the edition of provenance

2.2.3 Involvement of Lewistown residents' inspired poems

In Figure 23, the visualisation of the Mention-Network, which highlights poems inspired by real individuals from Lewistown, shows that most of these characters are part of the main connected component and are interconnected. Metrics for these nodes are presented in Table 14.

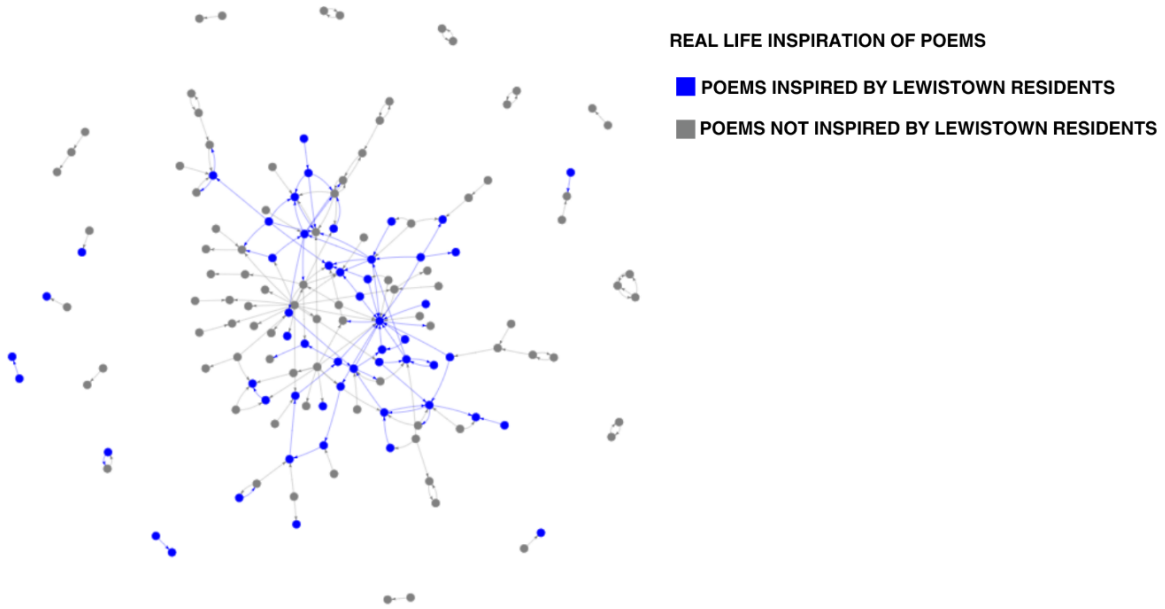

Fig. 23: Visualisation of the Mention-Network in the Spoon River Anthology with poems colored by real life inspiration of poems from Lewistown residents (in blue)

Metrics from poems inspired by Lewistown residents	**Value**
Number of poems present in the Mention-Network	53
Number of nodes present in the largest connected component	44

Tab. 14: Summary of metrics of poems inspired by Lewistown residents

The histogram showing the distribution of indegree for poems inspired by Lewistown residents and present in the largest connected component (Figure 24-A) indicates indegree values ranging from 0 to 17, with a gap between 7 and 16. The highest frequency, 14, corresponds to an indegree value of 1. The frequency is consistent at 6 for values 2 through 4, drops to 3 for a value of 5, and decreases further to 1 for values 6 and 17. For the same set of poems, the histogram for outdegree distribution (Figure 24-B) reveals values ranging from 0 to 6, with the maximum frequency of 17 at an outdegree of 1. The frequency then decreases to 8 for value 2, and to 4 for value 3, with frequencies of 2 and 3 for higher outdegree values of 4 and 5, respectively. Metrics for the indegree and outdegree values of this sample are compiled in Table 15.

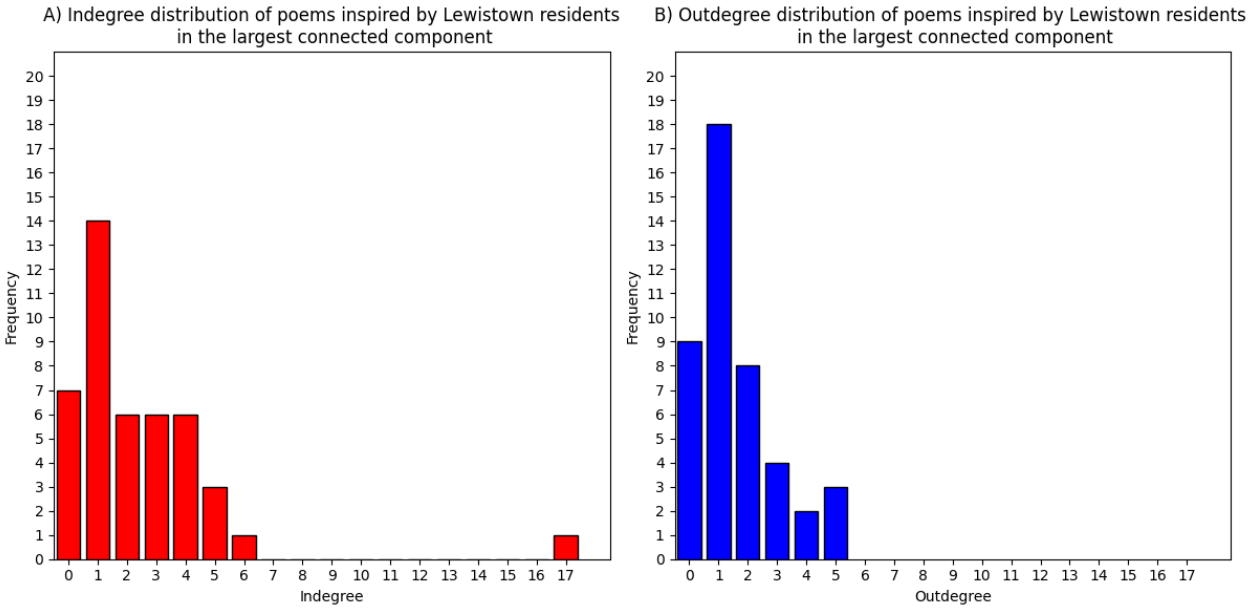

Fig. 24 : Histograms representing (A) the indegree distribution and (B) the outdegree distribution of poems inspired by Lewistown residents and belonging to the largest connected component

Values from indegree and outdegree of nodes of poems inspired by Lewistown Residents	**Value**
Number of nodes in the largest connected component with indegree above the average	23
Number of nodes in the largest connected component with indegree above the average	17
Maximum indegree of nodes	17
Maximum outdegree of nodes	5

Tab. 15: Summary of values extracted from the indegree and outdegree of nodes of poems inspired by Lewistown Residents

The histogram showing the PageRank distribution for the sample of nodes (Figure 25-A) covers values from 0.01 to 0.08, with values from 0.00 to 0.01 omitted for clarity. The peak frequency of 7 occurs in the lowest range, decreasing to 5 for the range 0.02 to 0.03. The frequency is steady at 1 for the ranges 0.04 to 0.06 and 0.07 to 0.08, with gaps between 0.03 and 0.04 and between 0.06 and 0.07. The histogram for betweenness centrality (Figure 25-B) shows values from 0.01 to 0.04, with a peak frequency of 6 for the range 0.01 to 0.02. Values between 0.00 and 0.01 were omitted for better visualisation. The frequency remains at 3 for the ranges 0.02 to 0.04. Metrics for PageRank and Betweenness Centrality are summarized in Table 16.

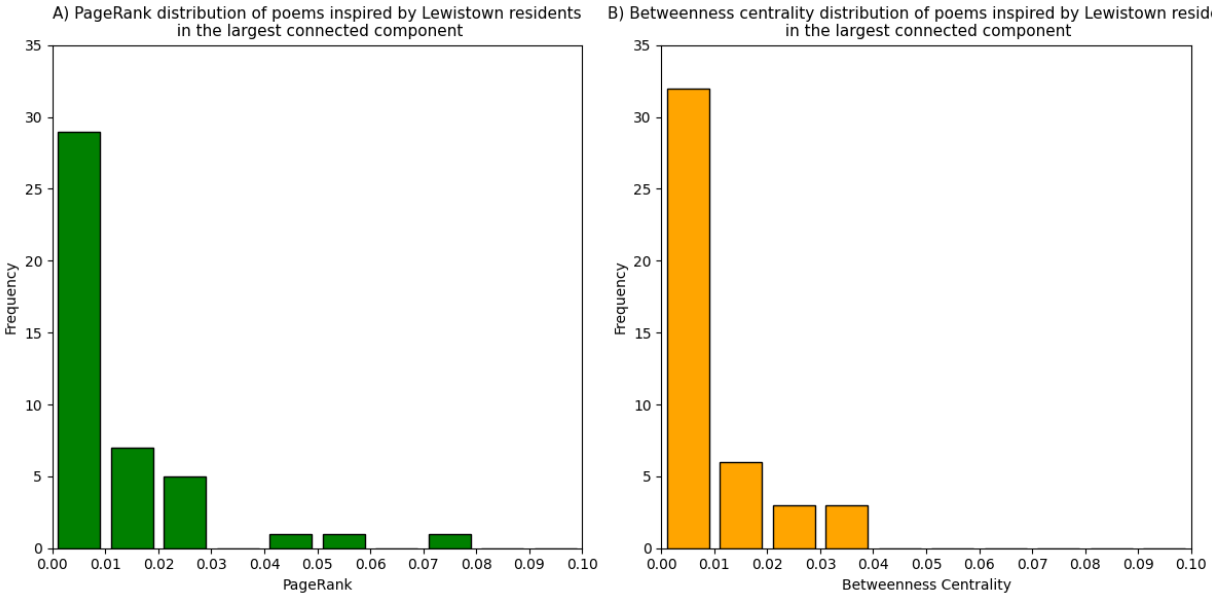

Fig. 25: Histograms representing (A) the PageRank distribution and (B) the Betweenness Centrality distribution of poems inspired by Lewistown residents and belonging to the largest connected component

Metrics from PageRank and Betweenness centrality of nodes of poems inspired by Lewistown Residents	**Value**
Number of nodes in the largest connected component with PageRank above the average	13
Number of nodes in the largest connected component with Betweenness Centrality above the average	13
Maximum PageRank value of nodes	0.076
Maximum Betweenness Centrality value of nodes	0.039

Tab. 16: Summary of values extracted from the PageRank and Betweenness Centrality of nodes of poems inspired by Lewistown Residents

The visualisation of the network of nodes representing poems inspired by Lewistown residents in the largest connected component (Figure 26) shows two connected components, with one encompassing nearly all the nodes. Metrics for this network are detailed in Table 17.

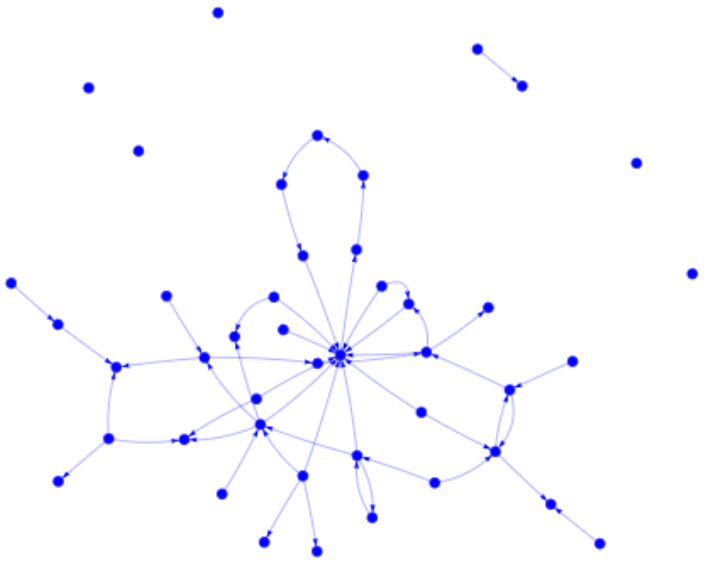

Fig. 26: Visualisation of the mention-network formed by the nodes of the poems inspired by Lewistown residents inside the largest connected component of the Mention-Network with spring length augmented

Metrics from the nodes present in the largest connected component of the Mention-Network and inspired by Lewistown residents	**Value**
Number of nodes in the largest connected component	37
Number of edges in the largest connected component	49
Density	0.0368
Average clustering coefficient	0.0297

Tab. 17: Summary of metrics from the largest connected component of the Mention-Network with only the poems inspired by Lewistown residents present

The analysis of the largest connected component identified six interconnected "Hub and Spokes" structures. Details of these structures and their metrics are provided in Table 18, with their visual representation shown in Figure 27.

Metrics of the "Hub and Spokes" structures in the Mention-Network made by poems inspired by Lewistown residents	**Value**
Number of Hubs	6
Number of shared spokes between structures	4
Number of not shared spokes	11
Average nodes in the structures	5.5
Average edges in the structures	5.7
Number of nodes in the smallest structure	5
Number of edges in the smallest structure	4
Number of nodes in the biggest structure	8
Number of edges in the biggest structure	9
Number of nodes in total	21

Tab. 18: Summary of metrics from the "Hub and spokes" structures found inside the largest connected component composed by poems inspired from Lewistown residents inside the Mention-Network

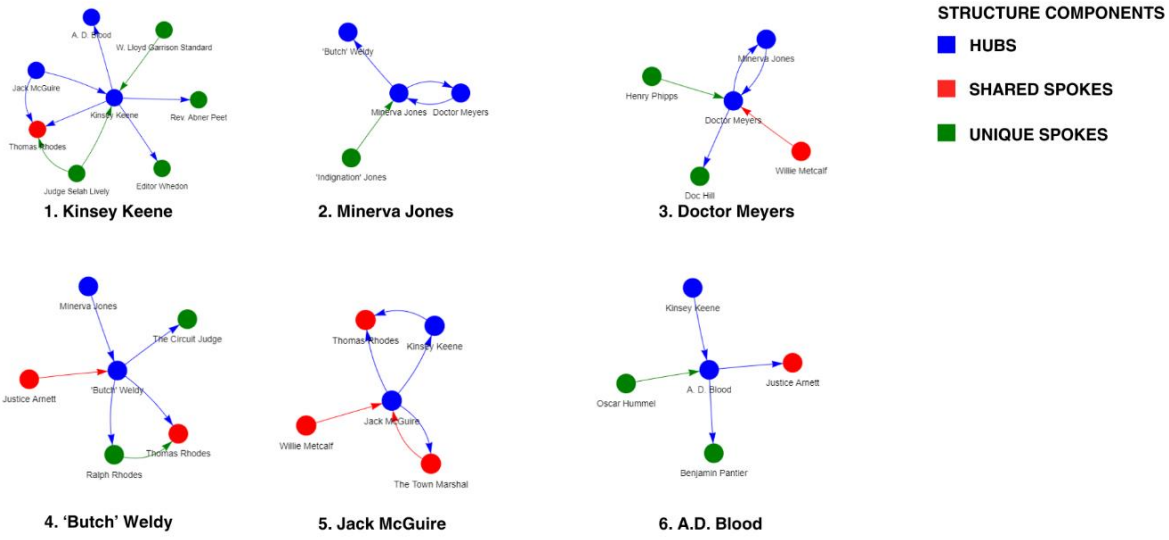

Fig. 27: Hub-and-spoke structures within the mention-network's largest connected component composed by poems inspired from Lewistown residents

The visualisation of the connections between hubs reveals two distinct connected components: triadic types 021C and 111U (Fig. 28).

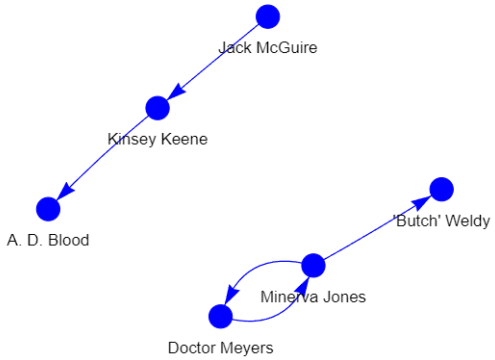

Fig. 28: Network Visualisation of the Hubs of the "Hub and Spokes" structures in the mention-network of Spoon River Anthology considering only poems inspired by Lewistown residents

The visualisation of the network, focusing on shared spokes and hubs (Figure 30), illustrates how these shared spokes integrate the triads into a single connected component. A triangular triad of type 030T is evident.

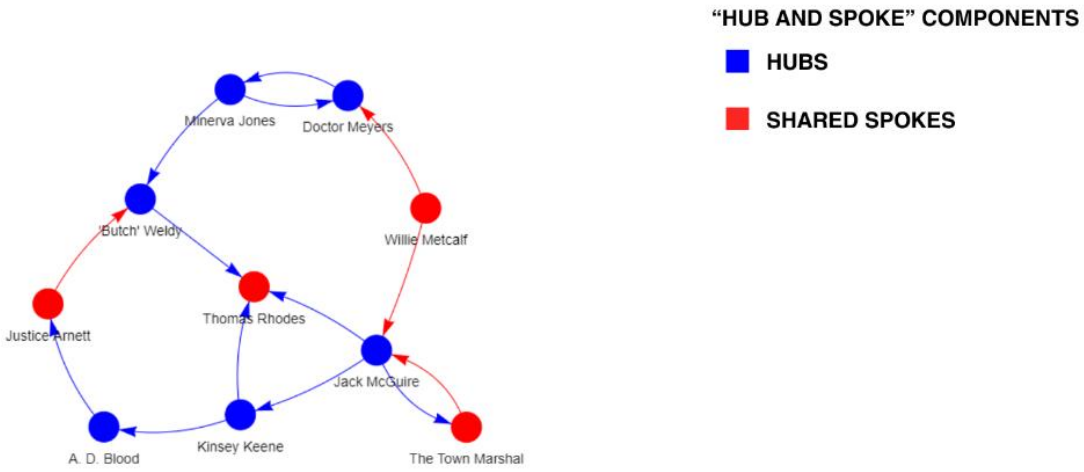

Fig. 29: Network Visualisation of the Hubs (in blue) and of shared spokes (in red) of the "Hub and Spokes" structures in the mention-network of Spoon River Anthology considering only poems inspired by Lewistown residents

Considering the connections with non-shared spokes, two triangular triads of type 030T are observed (Figure 30).

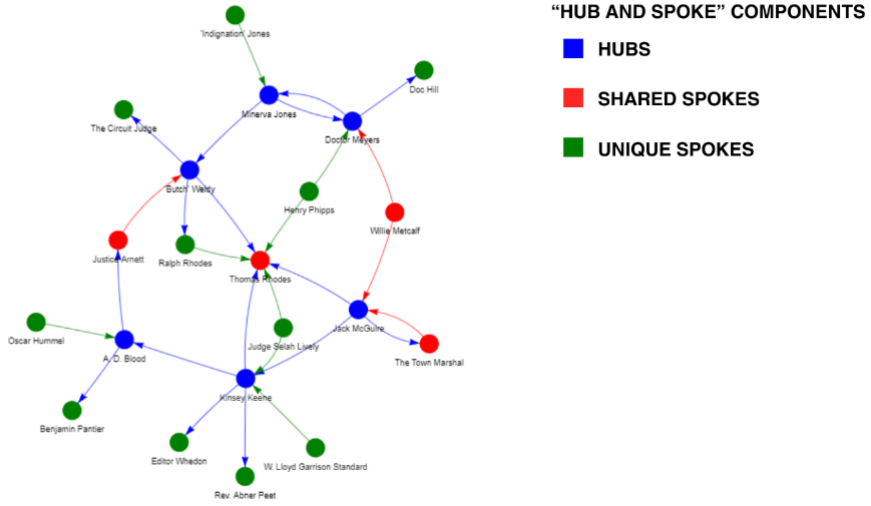

Fig. 30: Network Visualisation of the hubs (in blue), shared spokes (in red) and unique spokes (in green) of the "Hub and Spokes" structures in the mention-network of Spoon River Anthology considering only poems inspired by Lewistown residents

Figure 31 illustrates the arrangement of the "Hub and Spokes" structures within the connected component formed by nodes of poems inspired by Lewistown residents. It also shows the formation of a final triangular triad of type 030T. A cycle is created through a shared spoke connecting nodes outside the "Hub and Spokes" structures.

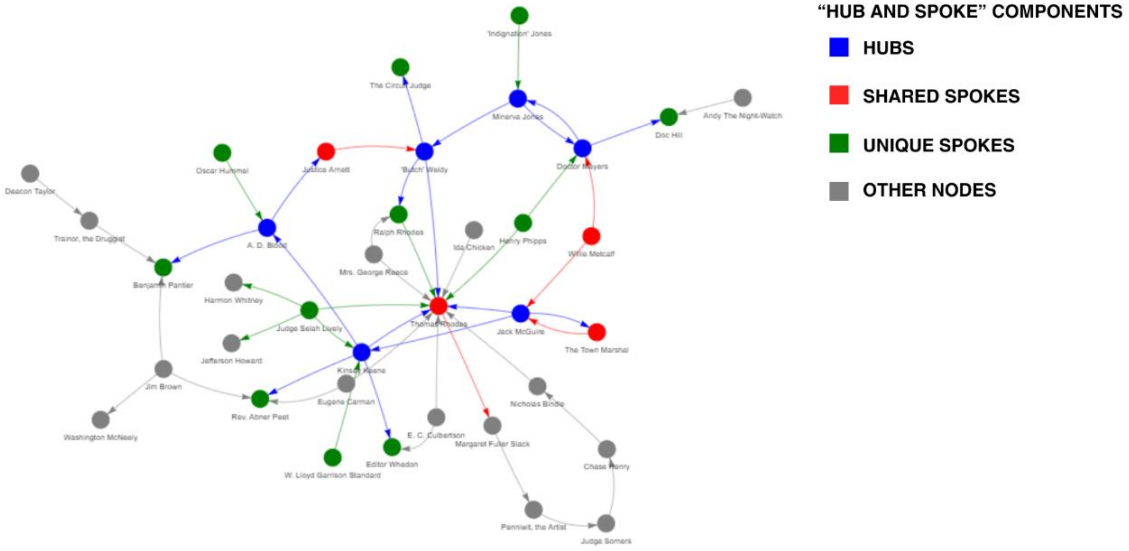

Fig. 31: Network Visualisation of the hubs (in blue), shared spokes (in red), unique spokes (in green) and other nodes (in gray) of the "Hub and Spokes" structures in the mention-network of Spoon River Anthology with only poems inspired by Lewistown residents

2.2.4 Sentiment of interactions with Lewistown-Inspired Poems

The data collected from the sentiment analysis of mentions directed to characters in poems inspired by residents of Lewistown are detailed in Table 19.

Metrics on edges towards Lewistown residents and their Sentiment Analysis	Value
Number of edges towards nodes of poems inspired by Lewistown residents	112
Edges with Positive (1) values of Sentiment	7
Edges with Neutral (0) values of Sentiment	52
Edges with Neutral Negative (-1) of Sentiment	53
Number of Nodes of poems by Lewistown residents with Negative Sentiments towards similar nodes	30
Number of Nodes of non-Lewistown poems with Negative Sentiments towards Lewistown-inspired nodes	22
Number of corrected data for Sentiment Analysis made by Chat Gpt-4o mini	11

Tab. 19: Summary of data obtained from the analysis of edges towards nodes of poems inspired by Lewistown Residents and their Sentiment

Figure 32 shows the Mention-Network with edges colored and sized according to sentiment. Most nodes connected by negative sentiment edges are closely clustered within the network.

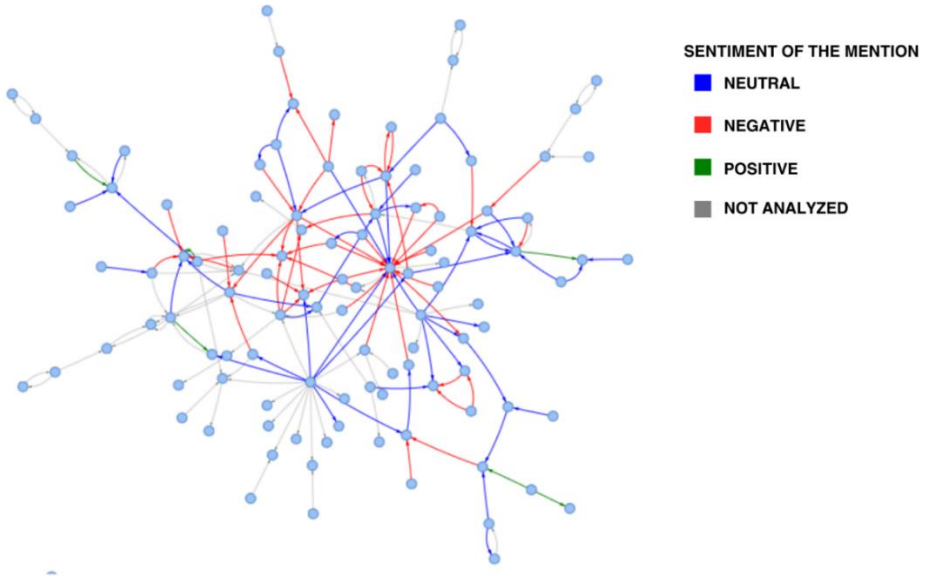

Fig. 32: Visualisation of the largest connected component of the Mention-Network with edges colored by sentiment detected in the Sentiment Analysis

The histogram of indegree values for nodes based on edges with negative sentiment (Figure 33) ranges from 0 to 12, with a gap between 6 and 11. The peak frequency is 12 at an indegree of 1, followed by a notable decrease of 7, with frequencies remaining constant at 1 for values greater than 3, indicating a right-skewed distribution.

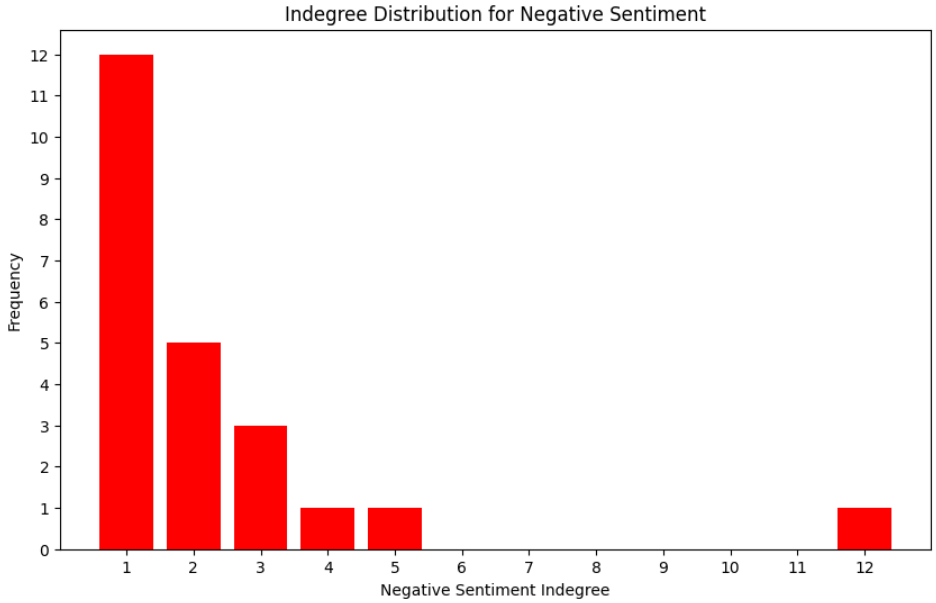

Fig. 33: Histogram representing the distribution of Negative Sentiment Indegree values with the corresponding node frequencies.

Figure 34 displays the adjacency matrix with colors representing edge attributes from sentiment analysis. It highlights a vertical neutral pattern throughout the Source Poems extending to Target Poems in the 10-30 range and a horizontal negative pattern from the Source Poems within the same range, ending before the 130th poem. Most negative sentiment mentions are directed towards poems in the 10-30, 40-80, and 90-110 ranges.

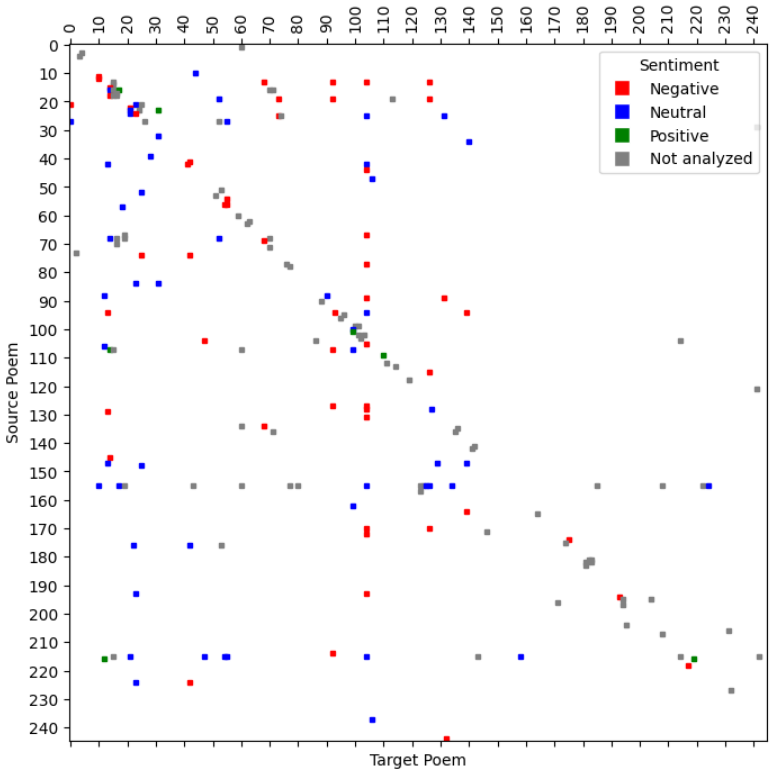

Fig. 34: Adjacency matrix of the Mention Network of the Spoon River Anthology, with axes ordered based on the appearance of the poem and interaction colored according to Sentiment Analysis value of the mention

2.3 Limitations, mitigating measures, and strengths of the analysis

This study acknowledged a few limitations, which have been systematically addressed through potential solutions and mitigating measures. One primary limitation is the methodology's applicability due to the unique structure of the Anthology. To address this, alternative approaches, such as analysing collections of interconnected texts similar to citation networks, may offer better insights. The stages of close reading, manual data extraction, and collection are prone to human error and can be time-consuming. To minimize these issues, the data collection process was repeated several times, and standardized methodologies for extraction and recognition were rigorously applied, as detailed in the extraction methodology subsection. Indirect mentions that could not be definitively attributed to specific characters were excluded from the database to improve accuracy. For sentiment analysis, a semi-automated approach using ChatGPT-4o-mini was chosen to balance efficiency with precision while maintaining necessary manual oversight. The metrics and visualisations were applied in line with the methodologies

outlined in Labatut and Bost's 2019 survey, ensuring that the approaches used were reliable and well-established for similar research. The strengths of this research are reflected in the effective integration of quantitative methods, which significantly reduced analysis time and enhanced its effectiveness. Although the initial time investment for data collection was substantial, the efficiency gained through quantitative methods more than compensated for this effort. The combination of qualitative and quantitative approaches facilitated a thorough analysis of the Anthology, an undertaking previously unexplored due to the work's complexity and fragmentary nature. This research fills gaps in the existing literature by applying Character Network Analysis in a new context, advancing both theoretical and practical knowledge of the methodology, and providing deeper insights into the author and the work. The interdisciplinary approach successfully integrates diverse fields, merging their methodologies and introducing new qualitative analysis methods that benefit both digital practitioners and experienced scholars.

3.4 Discussion

The analysis of the interpoetic narrative within the work reveals how Masters integrated this element into his poetry collection. Starting with the observation that Masters connects the stories mainly through references to the poems' titles, locating the characters becomes straightforward. This would explain why editions of the anthology have consistently included an alphabetical index rather than one arranged by the order of appearance of the poems, starting with the first edition. A key finding is that approximately 60% of the poems (144 out of 244) participate in the interpoetic narrative, interacting with one another, which highlights the importance of character interaction within the Anthology (*Tab.1*). It is noteworthy that these interactions do not encompass all the poems involved. The identification of 18 distinct connected components suggests that the goal of creating a fully interconnected microcosm within the Anthology was not entirely achieved (*Fig. 6*). Nevertheless, a significant proportion of the interacting poems (107 out of 144, or 74%) are situated within a single connected component, indicating a level of narrative cohesion (*Tab. 2*). Yet, the low interaction density within this component points to limited connectivity and fewer complex interactions between the poems, as evidenced by the low average clustering coefficient (*Tab. 2*). Further analysis reveals that the majority of poems are either infrequently mentioned or not mentioned at all, whereas

33% of the poems exhibit above-average values for both indegree and outdegree (*Fig. 6; Tab. 3*). This suggests that certain poems play a more prominent role within the interpoetic narrative, either by being frequently mentioned or by frequently mentioning others, particularly in relation to characters with significantly above-average values. The examination of PageRank and Betweenness centrality metrics clarifies this pattern. Although a subset of frequently cited characters exists, they are not necessarily interconnected. Only a small number of poems are both central within the network and interconnected. Specifically, although 25 characters exhibit above-average PageRank values, only 4 characters are considered both highly mentioned and interconnected (*Fig. 7; Tab. 4*). This implies that the narrative is largely driven by characters who, while mentioning others, are themselves infrequently mentioned, indicating a reliance on peripheral figures rather than central ones. Moreover, the potential for interweaving poems and characters is limited, as shown by the distribution of Betweenness centrality. This metric indicates a scarcity of characters acting as critical connectors between multiple stories or narratives, suggesting a predominantly linear narrative structure. The absence of characters whose removal would significantly disrupt the mention interaction network points to a lack of deeply interconnected perspectives across groups of characters (*Fig. 7; Tab. 4*). When discussing this as a narrative device, it is worth noting that the interpoetic narrative seems to aim to connect stories while maintaining an underlying layer of mystery. This is achieved by the deliberate lack of connections among influential characters and the progression of the narrative through a few key figures who serve as bridges between pairs of poems. Regarding the placement in the text, Masters's organization of the interpoetic narrative in the literary work appears to follow a deliberate strategy, primarily relying on direct mentions, often through the poem's title, to establish connections between poems. A considerable number of semi-direct and indirect mentions are also observed. A noticeable pattern emerges whereby similar types of mentions are frequently reused for the same character, with Masters experimenting with a mixed approach in only a small subset of poems (*Tab. 5*). The organization of the narrative is further characterized by a tendency to create continuity by positioning many interactions close to each other, as shown by the primary diagonal of the adjacency matrix, particularly in three distinct sections of the book. The initial cluster of interactions, located at the beginning of the work, introduces characters who later appear in other major clusters, all situated within the first half of the Anthology (*Fig. 8*). This narrative strategy focuses

predominantly on the retrospective development of characters following their introduction, rather than on foreshadowing future characters and events, an element that appears only in the first poems. As a result, a noticeable decline in interactions among poems is observed in the latter half of the collection. The type of mention used is also correlated with the poem's position within the work (*Tab. 6*). For example, semi-direct mentions are predominantly associated with characters occupying specific positions within the collection (first half), while the use of indirect mentions aligns with the proximity of poems to one another. From a topological perspective for smaller connected components, Masters favors simple interactions between characters, typically forming dyads that interact either reciprocally or unidirectionally. Although more complex structures, such as a triad 210, are present, they are less prominent (*Tab. 7; Fig. 9*). The interaction structure at the center of the primary connected component within the Mention-Network is particularly notable, as it is dominated by Hub-and-Spoke configurations (*Tab.8; Fig. 10*). These hubs are interconnected, with both shared and unique spokes radiating from them (*Fig. 11*). The prevalence of these structures suggests that the narrative revolves around key characters whose stories are fragmented and explored through the experiences of numerous other characters. These Hub-and-Spoke structures indicate that many central characters play a crucial role in advancing the narrative and maintaining its continuity, as the narrative frequently returns to these characters or other key figures. Shared spokes function as bridges between different structures and characters, while unique spokes highlight characters to provide more perspectives on interactions, often through triad 030T (*Fig. 12; Fig. 13; Fig. 15*). This triad captures essential interaction dynamics, extending beyond hierarchical structures, in which characters narrate the actions or experiences of a single interaction between poems. Unfortunately, a predominantly linear connection remains, as 57 of the nodes belong to the hierarchical structures previously discussed (*Tab. 8*). This creates a significant gap in continuity, emerging from the midpoint of the Anthology onward, which can be observed both in the matrix and by directly reading the poems (*Fig. 8*). Representing the narrative structure graphically, we have something like what is seen in Figure 35.

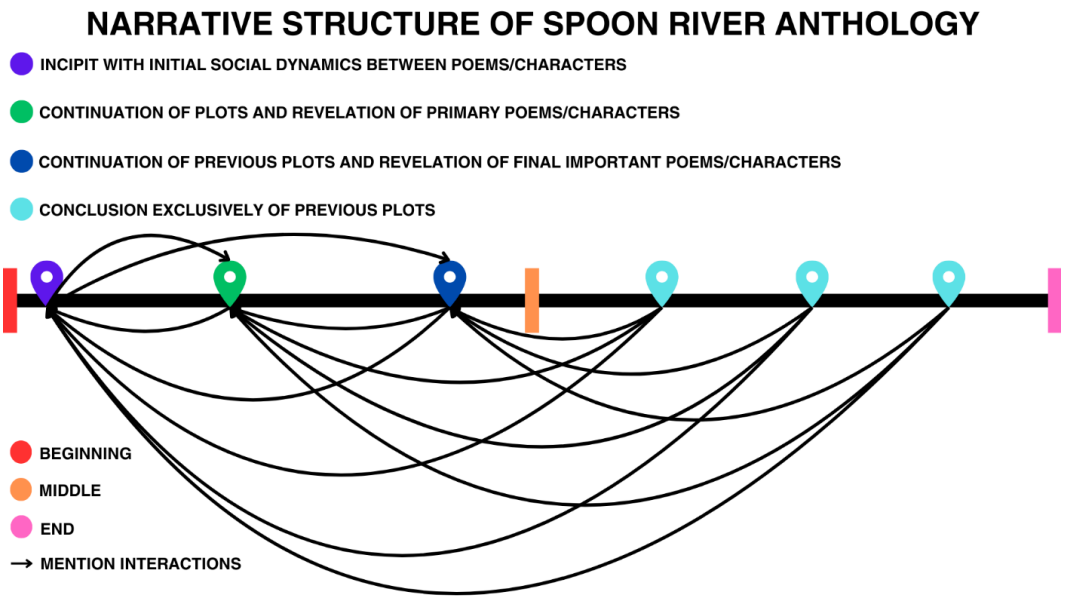

Fig. 35: Illustration of the narrative structure of *Spoon River Anthology* using a linear timeline

Reaccessing the text of this group of structures, reveals one of the central themes of the Anthology through the interpoetic narrative: the profound impact that words and opinions have on an individual's life, especially within a small community. This theme is evident in poems like *Aner Clute*, where the following passage highlights this idea:

"[...]Suppose a boy steals an apple
From the tray at the grocery store,
And they all begin to call him a thief,
The editor, minister, judge, and all the people-
'A thief,' 'a thief,' 'a thief,' wherever he goes.
And he can't get work, and he can't get bread
Without stealing it, why, the boy will steal.
It's the way the people regard the theft of the apple
That makes the boy what he is."

Similarly, in *Fiddler Jones*, the character reflects on societal expectations:

"[...]And if the people find you can fiddle,
Why, fiddle you must, for all your life.[...]"

But contrary to *Aner Clute* he ends with:

"*[…]I ended up with forty acres;*
I ended up with a broken fiddle-
And a broken laugh, and a thousand memories,
And not a single regret."

In small communities, gossip is widespread and has a strong impact on how individuals are perceived. For instance, *Aner Clute*, despite owning his epitaph and trying to assert his identity, finds his words and judgments influenced by societal views until his death. Similarly, in *Fiddler Jones*, although the character is a landowner, he is mainly remembered for his fiddling talent. He accepts this without feeling constrained by the label. Masters uses these characters to examine how others' perceptions shape one's self-image. This dual perspective aligns with the philosophical ideas of Levinas and Sartre. Levinas argues that the encounter with the Other and their gaze contributes to one's sense of self, while Sartre suggests that this gaze distances one from their own subjectivity, bordering on objectification (*Pimentel, 2023, pp. 4-5*). This theme is evident in the Mention Network, where many characters with high indegree values appear. The triads they form illustrate how an individual, despite attempts to express their self-concept through their epitaph, is often overshadowed by mentions from surrounding characters. This dynamic is reinforced by the central structures of the network, designed to sustain this pattern. To further this interpretation, examining the poems through these structures reveals a significant Hub-and-Spoke configuration, such as that of *Daisy Fraser*. Analysis shows that Daisy's influence, reflected in her three mentions, is mainly confined to the village's negative views on her sexual freedom. Her capacity to change or resist the circumstances reflected in her epitaph is limited. Masters suggests that characters' influence is shaped by others' judgments and perceptions within the anthology. This effect mirrors the real-world impact of words, as seen in *Aner Clute* and *Fiddler Jones*, where the poems influence characters' actions and experiences as though they held tangible power. Another example is Thomas Rhodes, the character with the most mentions. Network analysis reveals that very few of the frequently mentioned characters are directly connected. This lack of direct connection extends into a broader narrative where other characters indirectly highlight instances of social inequality perpetuated by figures who conspired together. Although these characters may not have faced consequences during their lives, they face repercussions in their poetic testament. In death, many characters denounce Rhodes's

actions, which were unchallenged during his life. This collective denunciation affects Rhodes in his own poem, where he is tormented by guilt for his pursuit of wealth and power, rather than enjoying a peaceful "eternity." Through this narrative device, Masters restores a moral balance, showing how perceptions and judgments of others can reshape and redefine an individual's influence. This approach reflects modern sensibilities, despite Masters' claims of avoiding contemporary literary movements (*Jewett, 1942, p. 37*). This perspective leads us to view the afterlife created by Masters as significant and interconnected with the earthly realm. In this afterlife, those who suffered and could not resist find their dignity restored, while those who caused harm to face their actions. The inclusion of characters with dubious reputations presenting their own versions of events, likely influenced by Masters' legal background, is integral to the work. Masters does not invent this tradition; rather, the influence of Greek literature in his work is evident in the interaction network and its matrix. Reflecting on authors like Xenophon[2], we recall the essay in Italo Calvino's *Perché leggere i classici*, where Calvino analyses Xenophon's *Anabasis* and discusses the pragmatic and ethical aspects reflected in the text. Calvino admires Xenophon, who, while leading his soldiers, is aware of the consequences of his actions driven by choice, anticipating a determined response from his enemies (*Calvino, 2009, pp. 27-28*). This parallels Masters' effort to present various perspectives while maintaining empathy, even for those responsible for the events depicted in the Anthology. Masters displays profound humanity, clearly conveying facts through the many voices that denounce the events and, in many cases, could only endure. It is difficult not to feel sympathy considering that Masters fell ill after submitting the Anthology to the publisher. By making leaps of imagination, this collective, which narrates and denounces the activities of only a few individuals, revealing things we would otherwise not have known and taking a clear stance, may resemble the role of the collective character of the chorus in classical tragedy, although in a fragmented and dispersed form within the Anthology. It gradually presents pieces that complete the puzzle of the main events described. However, it is right to leave these considerations to the humanists, which leads to showing the need for this type of analysis or projects to be made up of different actors from different disciplines.

[2] The Greek philosopher and historian from the 5th-4th century BCE is referenced by Masters in his 1920 Domesday Book in the poem titled *Father Whimsett*.

Turning to the 1916 edition, which incorporates 33 additional characters, it is worth noting that it did not significantly advance the interpoetic narrative, despite 15 of these new characters being integrated into it (*Fig. 17; Tab. 11*). Masters included characters mentioned slightly more frequently than average (2 above the average indegree). Most of the new characters consist of nodes with exceptionally high mention counts, including five characters with values well above the average (*Fig. 18; Tab. 12*). Some of these show elevated PageRank and Betweenness Centrality values, with one poem having the highest Betweenness Centrality (*Fig. 19; Tab. 13*). This suggests that Masters introduced fewer than five poems that play a substantial role within the narrative. Masters chose to include interactions and characters that extend the main narrative established by the three clusters of interactions around the primary diagonal of the adjacency matrix. Thus, the new edition mainly reinforced the existing trajectory of the Anthology rather than introducing a significant shift in the interpoetic narrative (*Fig. 20*). Nevertheless, there is an important point. By revisiting the text of *Kinsey Keene*, the character with the highest betweenness centrality and a hub within one of the interconnected structures, we can grasp his importance in the final edition. Kinsey Keene not only has a high indegree but also contributes early to the Anthology by identifying the antagonists. He is among the few characters who clarify the unfolding events, even though his portrayal is constructed entirely through other characters' perspectives. His words are direct and pointed, almost like an accusation, and through his presence, he offers a final perspective akin to that of a lawyer, much like the author himself. This perspective anticipates the unfortunate fate of those he mentions, and it is noteworthy that he is situated within the first cluster of poems along the primary diagonal. Analysing poems with the highest betweenness centrality in the new edition reveals significant insights into both this addition and the network itself, offering a perspective similiar to that obtained from studying poems with high PageRank. For instance, looking at characters with the highest betweenness centrality shows that many are used by Masters to highlight the harsh realities of rural Midwest life, such as Minerva Jones, 'Butch' Weldy, Blind Jack, and Jack McGuire, who hold prominent positions in the network. Conversely, characters with high PageRank, including Thomas Rhodes, Benjamin Pantier, Mrs. Benjamin Pantier, and Washington McNeely, represent affluent lives characterized by industrial and banking wealth, along with economic stability from extensive landholdings. This contrast suggests that class conflict and struggle within the literary work are also reflected in the mentions.

Concerning the poems inspired by the residents of Lewistown, several observations address many existing questions. Notably, there is a significant number of such poems within the network, positioned closely together, suggesting that Masters intentionally facilitated interactions among these poems (*Fig. 21; Tab. 14*). This is further supported by their presence as the majority of those with above-average values in key metrics such as indegree, outdegree, PageRank, and Betweenness Centrality. Specifically, these poems display the highest indegree values, and the highest PageRank and Betweenness Centrality values, indicating their crucial role in interacting with other characters, both in terms of receiving and making mentions, and in linking pairs of poems (*Fig. 22; Tab. 15; Fig. 23; Tab. 16)*. Furthermore, most of these poems interact directly, though they do not form a particularly dense or cohesive network, as reflected by the low density and average clustering coefficient values. The similarity of this network to the overall Mention Network suggests that Masters may have used this network of stories inspired by real individuals as a foundational basis for constructing the rest of the Anthology (*Fig. 6; Tab. 2; Fig. 24; Tab. 17*). Similarly, an analogous interaction pattern was observed in a study examining student interactions in an online asynchronous discussion forum (*Yusof & Rahman, 2009, p. 28*). The interaction patterns among students in that study resemble the interaction network of mentions among characters in the Anthology inspired by real individuals (*Fig. 25-28*), despite the directional graph in the study being weighted. An analogous adjacency matrix (*Fig. 32*), similar to the one seen in the first half of the Mention-Network's poems in the Anthology, was found in individual discussions on online forums, as described in the study by Varcellone-Smith, Jablokow, and Friedel (*Varcellone-Smith et al. 2012, p. 227*). For now, this correlation remains intriguing rather than conclusive. The situation grows more complex when considering the results of sentiment analysis within the mention network (*Tab. 19*). While ChatGPT-4o mini was effective in most cases analysed, the sentiment directed at characters inspired by Lewistown residents tends predominantly toward neutral and negative (*Fig. 30*). Notably, 56% of these negative sentiments come from characters also inspired by Lewistown residents, with the rest attributed to characters presumably created by Masters. This uneven distribution of negative sentiment, particularly within interacting groups of characters, suggests a distinct cluster of poems marked by negative or neutral sentiments. Many of these sentiments are tied to characters deliberately created by Masters. This pattern reflects the author’s critical perspective on Lewistown and suggests that this view was not

subtly hidden, helping to explain the strong backlash and subsequent censorship of the book. Moreover, despite the presence of specific characters who are primary targets of negative sentiment, the sentiment is widely distributed across numerous characters, albeit with varying intensities (*Fig. 31*). The placement of the poems within the Anthology does not significantly obscure this distribution, as negative mentions are concentrated within the three main clusters around which the interpoetic narrative is structured (*Fig. 32*). Masters not only presented a distinct version of rural life but also revealed it clearly through the language used. This observation invites further consideration. Masters, as a lawyer, would have been acutely aware of the ethical implications of his work. Given this, the structure and content of his work might have affected his career both as a poet and a legal professional. Despite this awareness, Masters included recognizable stories and characters, especially from local contexts, as central elements of his Anthology. It is well-documented that his transition from Petersburg to Lewistown was a significant turning point influenced by his experiences and his parents' perspectives. Nevertheless, one might question why Masters did not attempt to obscure or justify his intentions, as seen in Alessandro Manzoni's use of a fictional pretext in *I Promessi Sposi* to critique contemporary issues under the guise of historical events (*Sorrentino, 2020, pp. 339-340*). Masters appears to have done the opposite, raising questions about his rationale. One certainty is that the use of a pseudonym did little to hide the author's identity, especially considering how quickly he chose to reveal himself. In retrospect, maintaining the pseudonym "Webster Ford" for a longer period and setting the work 30 years earlier might not have significantly altered its impact.

CONCLUSIONS

This thesis investigated the use of Character Network Analysis within the field of narratology, building on a preliminary exploration. The discipline of narratology evolved, especially with its shift toward the concept of agency, and began employing digital methods to examine works of fiction through character interactions. Significant effort was dedicated to addressing recent themes and shared values relevant to Digital Humanities, as this research fell within an interdisciplinary framework. Consequently, it adopted a mixed-methods approach that integrated both qualitative and quantitative analyses. A meta-research framework based on CNA was developed to facilitate the application and study of this analytical methodology. This framework aimed to benefit scholars in both computer science and the humanities by focusing on gaps in the application and understanding of the methodology. It utilized a work of fiction that required this analytical support. The research successfully addressed the research questions of both studies conducted. The study examined a neglected aspect of the Spoon River Anthology by employing quantitative methods to uncover how the author interconnected the various stories of the fictional town in late 19th-century Illinois. One of the main findings highlighted the importance of interpoetic narration within the Anthology, as it comprises more than half of the poems, with 74% found in a single interconnected component. This demonstrates a significant effort to weave together a multitude of stories and characters. Among these, only a few are very important, either through frequent mentions or as frequent mentioners. Furthermore, only a limited number of characters are both central and interconnected, forming a non-linear narrative with Hub-and-Spoke structures. It is through these structures that we learn about important characters and their actions. The connections between them render the narrative concise while being rich in perspectives and relationships. This pattern illustrates how the fragmented stories of central characters are woven together through their interactions. The study also suggests that the poem arrangement within the Anthology could have been improved. While circular interactions among characters are quickly abandoned for retrospective developments, this shortcoming is partly addressed by examining the average distance between mentions, revealing a logic behind the positioning of the mentions. The new edition's impact did not significantly alter the narrative direction but instead reinforced it, with characters like Kinsey Keene adding crucial perspectives. The findings show that

Character Network Analysis can be highly effective for poem collections or similar text-based content. The interconnected structure of the Spoon River Anthology makes it particularly suitable for this form of analysis. Future research could extend this method to other collections of texts where similar relational analyses might offer insights. Regarding the use of the network to study the poems again for new insights, it is important to note that the visual cues provided a fascinating opportunity to reevaluate many internal structures of the network, including those related to characters and poems. This close reading facilitated an exploration of social interactions and gossip, which are recurring themes in the work. In the discussion, a shift in the agency of the characters was observed based on the mentions they received from other poems. This demonstrates how societal perceptions shape identities, influencing self-conception and moral evaluation. The research also identified a notable division regarding class struggles, illustrated through betweenness centrality and PageRank, which could indicate a foundation of empathy toward the oppressed and an inability to effect change due to those who possess the means of production. Regarding the semi-automatic analysis, which incorporated both quantitative and qualitative methods, it revealed an uneven distribution of negative sentiment, primarily directed toward characters based on Lewistown residents. This distribution reflects Masters' critical view of his hometown, which contributed to the work's censorship and reception issues. The mention network, built upon interactions among Lewistown residents, further underscores the importance of societal and personal perspectives on literature. From the perspective of narratological study, CNA appears to be a useful tool when applied consciously and adapted to the chosen work of fiction. In fact, this analysis made it possible to confirm existing theories or suspicions regarding the author and the work, potentially reopening studies on the Anthology from a renewed perspective. It allowed for an observation of the narrative structure, both in terms of its spatial organization within the work and from the perspective of character construction, where roles or characteristics could be identified through the selected evaluation metrics. The ability to obtain visualisations for comparison in analysing the network with real data from similar situations yielded modest yet satisfactory results. In conclusion, the work proved valuable for analysing themes that might otherwise be difficult to address. Utilizing this new approach in various ways demonstrated that both qualitative and quantitative methods greatly benefited from the use of a Network Graph, offering a fresh perspective on the work. This suggests the great utility of creating a comprehensive digital

edition of the Anthology with interactive features demonstrated during this research. At present, no edition fully captures the textual fluidity of the work or offers easy access for in-depth study, which could be a goal for future developments. In this context, I reached out to Masters scholars to evaluate the need for a digital edition with the features studied, to see if there was also academic interest in these topics. Through this research, I was able to get in touch with Julianne VanWagenen and Jason Stacy, both of whom supported the idea of my project and the possibility of visualizing character interactions within the Anthology for its study. Further discussions with VanWagenen emphasized the value of a tagging system for poems to enhance research usability; the ability to independently assign attributes would make users' work more autonomous and less dependent on digital expertise. Therefore, creating a digital edition with improved navigability of character interactions and implementing automatic visualisations through markup could offer significant scholarly benefits. In line with the principles of Digital Humanities, I have not only created a repository for the datasets used in this project but also included a "Read Me" section, where I have provided the necessary code to upload these datasets into a personal ipynb file and analyse them using the relevant libraries, which are also available in the repository (*Appendix A*). Supplementary materials, such as video tutorials, were provided to assist those less familiar with the analysis process. As part of the project's dissemination goals, an HTML page hosted on GitHub was developed, and can be found in *Appendix C*, which features a dynamic visualisation of the network through the integration of the PyVis file into the HTML as can be seen from Figure 36.

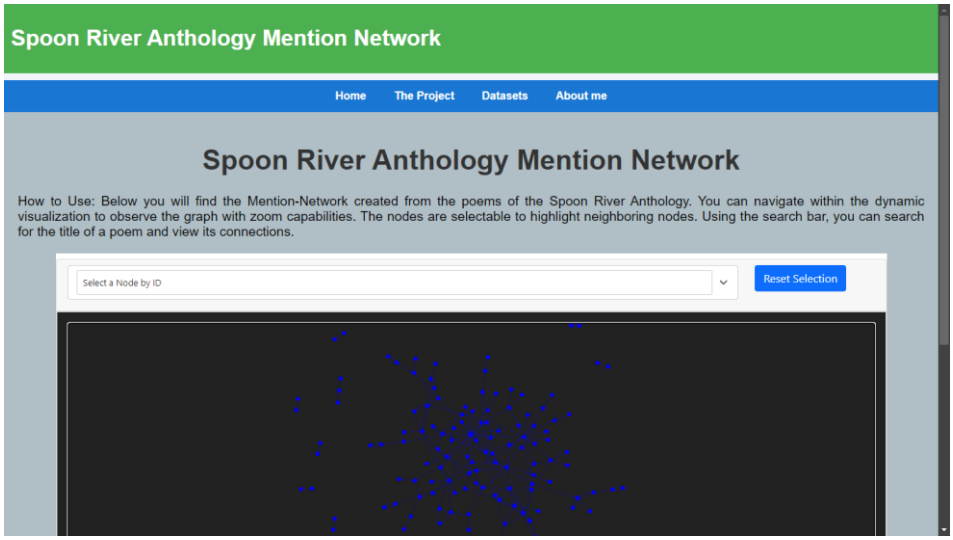

Fig. 36: Screenshot from the GitHub Pages site showcasing the home section of the Spoon River Anthology Mention Network website

This visualisation is equipped with a search bar for the titles of the poems. The navigation bar of the homepage further includes a page outlining the characteristics of my thesis project, a link to the GitHub repository for downloading the datasets, and an "About Me" page that provides a brief overview of my background. The two repositories on GitHub are made accessible under the Attribution-NonCommercial-ShareAlike 4.0 International licence, also referred to as LICENSE-CC-BY-NC-SA. Although this does not directly relate to the thesis itself, as this option is not available, it is essential for the research to remain aligned with its core objectives. Especially when using methodologies applicable to other fields, it is important to consider the involuntary contributions that could be made. Through this project, I have aimed to contribute to the growing intersection between computational methods and literary analysis, demonstrating how Digital Humanities can provide new pathways. The integration of digital tools with traditional literary approaches offers a promising avenue for future research, fostering more accessible and dynamic ways to engage with complex narratives and literary works in general. The productive and trusting collaboration between the humanities and the sciences appears to be increasingly within reach.

ACKNOWLEDGEMENTS

I would first like to thank my supervisor, Fabiana Zollo, who, despite her academic commitments, managed to dedicate time and provide me with her support. I also thank my co-supervisor, Franz Fischer, who, despite the short notice, agreed to take on this role and assist me during the final and most challenging stages of this work.

I would like to express my gratitude to scholars Julianne VanWagenen and Jason Stacy for their valuable assistance and the interest they showed in my project, despite the continental distance between us.

I am grateful to my parents, whose sacrifices have allowed me to pursue a continuous academic career in a place where I could finally discover my own identity and shape my future. I would like to greet and thank my siblings, who see me rarely, but whom I always think of. I also thank those friends and colleagues who, during these difficult years, lifted me up and accompanied me to the end, like a modern Dorando Pietri. To them, and to my remarkable ability to keep going, I owe everything.

And to all the children... Congratulations!

APPENDIX

Appendix A – Dataset created and utilised

GitHub repository containing the datasets: https://github.com/MMoscato00/Spoon-River-Mention-Network

Appendix B – Prompt and original codes written

B-1: Prompt provided to Chat GPT4o-mini for Sentiment Analysis

"*I will provide you with a text, where specific names of people are enclosed in brackets. For each name mentioned, analyse the sentiment associated with it in the context of the surrounding text. Provide the sentiment as 'Positive,' 'Negative,' or 'Neutral,' and assign a sentiment value on a scale from -1 (very negative) to +1 (very positive), with 0 being neutral. After each sentiment and value, give a very brief explanation (no more than one sentence) to justify your analysis.*
Example format for each analysis:
Name:[John]
Sentiment: Positive
Value: +0.7
Explanation: The text expresses admiration and respect towards John.
Here is the text: "

B-2: Python code written for extracting 'Hub and Spokes' structures according to metrics found

```
def find_hub_and_spoke_structures(graph):
    structures = []
    for node in graph.nodes:
        successors = list(graph.successors(node))
        predecessors = list(graph.predecessors(node))
```

```
        if len(successors) > 1 and len(predecessors) > 1:
            structure = [node] + successors + predecessors
            if structure not in structures:
                structures.append(structure)
    return structures

hub_and_spoke_structures = find_hub_and_spoke_structures(G)
print("Hub-and-Spoke structures in the graph:")
for structure in hub_and_spoke_structures:
    print(structure)
```

Appendix C – Dissemination materials

C-1: Project Web Page:

Link to the project webpage:
https://mmoscato00.github.io/MentionNetworkSpoonRiver/

C-2: Project Web Page HTML and CSS files:

Link to the repository: https://github.com/MMoscato00/MentionNetworkSpoonRiver

Appendix D - List of figures

Appendix E - List of tables

REFERENCES

BIBLIOGRAPHY

Agarwal, D., & Vijay, D. (2021, May). *Genre classification using character networks*. In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 216-222). IEEE.

Ahnert, R., Ahnert, S. E., Coleman, C. N., & Weingart, S. B. (2020). *The network turn: Changing perspectives in the humanities*. Cambridge University Press.

Al-Taie, M. Z., & Kadry, S. (2017). *Python for graph and network analysis* (pp. 1-184). Cham: Springer International Publishing.

Amalvy, A., Labatut, V., & Dufour, R. (2024). *Renard: A Modular Pipeline for Extracting Character Networks from Narrative Texts*. Journal of Open Source Software, *9*(98), 6574.

Amerian, M., & Jofi, L. (2015). *Key concepts and basic notes on narratology and narrative*. Scientific Journal of Review, 4(10), 182-192.

An, Y., Zhang, Y., & Zeng, B. (2015). *The reliable hub-and-spoke design problem: Models and algorithms*. Transportation Research Part B: Methodological, 77, 103-122.

Ardanuy, M. C., & Sporleder, C. (2015). *Clustering of novels represented as social networks*. Linguistic Issues in Language Technology, 12.

Barnstone W. (1968). *Introduction*. In Masters, E.L. (1968). *The New Spoon River*. Macmillan. XVII- XXVI

Beeferman, D., Berger, A., & Lafferty, J. (1999). *Statistical models for text segmentation*. Machine learning, *34*, 177-210.

Berhe, A., Guinaudeau, C., & Barras, C. (2023). *Survey on narrative structure: from linguistic theories to automatic extraction approaches*. Revue TAL: traitement automatique des langues, 63(1).

Bimantara, I. M. S., Purwitasari, D., Er, N. A. S., & Natha, P. G. S. (2024). *Balinese story texts dataset for narrative text analyses*. *Data in Brief*, *56*, 110781.

Boyd, R. L., Blackburn, K. G., & Pennebaker, J. W. (2020). *The narrative arc: Revealing core narrative structures through text analysis*. Science advances, 6(32), eaba2196.

Burgess, C. E. (1987). *" Spoon River": Politics and Poetry*. *Papers on Language and Literature*, *23*(3), 347.

Calvino, I. (2009). *Perché leggere i classici*. Edizioni Mondadori.

Carroll, W. (2021). *Small-town Styx: Edgar Lee Masters, Rural Confinement and Pastoral Subversion*. Comparative American Studies An International Journal, 18(1), 118-137.

Chandler, J. C. (1921). *The Spoon River Country*. Journal of the Illinois State Historical Society (1908-1984), 14(3/4), 249-329.

Chatman, S. (1975). *Towards a theory of narrative*. New literary history, 6(2), 295-318.

Chen, C. M., Chang, C., & Chen, Y. T. (2023). *A character social network relationship map tool to facilitate digital humanities research*. Library Hi Tech, 41(2), 516-542.

Chen, R. G., Chen, C. C., & Chen, C. M. (2019). *Unsupervised cluster analyses of character networks in fiction: Community structure and centrality*. Knowledge-Based Systems, 163, 800-810.

Ciğerci, F. M., & Yıldırım, M. (2024). *From Freytag pyramid story structure to digital storytelling: Adventures of pre-service teachers as story writers and digital story tellers*. Education and Information Technologies, 29(5), 5697-5720.

Ciotti, F. (2023). *L'analisi del testo*. In F. Ciotti (Ed.), Digital humanities. *Metodi, strumenti, saperi*. Carocci editore Spa.

Ciotti, F., Corradini, E., Cugliana, E., D'Agostino, G., Ferroni, L., Fischer, F., ... & Sahle, P. (2022). *Digital Scholarly Editions Manifesto*. UMANISTICA DIGITALE, 12, 103-108.

Cohn, N. (2013). *Visual narrative structure*. Cognitive science, 37(3), 413-452.

Dacos, M. (2011). *Manifeste des Digital humanities*. THATCamp Paris, 26.

Elkins, K., & Chun, J. (2019). *Can Sentiment Analysis Reveal Structure in a Plotless Novel?*. arXiv preprint arXiv:1910.01441.

Elli, T. (2023). *Supporting Literary Criticism with Data Visualisation: Four Design Guidelines for Facilitating Interdisciplinary Collaborations*. *CONVERGÊNCIAS*, *16*(32), 152-163.

Elson, D., Dames, N., & McKeown, K. (2010, July). *Extracting social networks from literary fiction*. In Proceedings of the 48th annual meeting of the association for computational linguistics (pp. 138-147).

Eve, M. P. (2022). *The digital humanities and literary studies* (p. 208). Oxford University Press.

Falcetto, B. (2019). *Vedere i testi, leggere le immagini: le visualizzazioni per l'"Atlante Calvino"*. In Tuttestorie di donne (pp. 203-206). il Saggiatore-Fondazione Arnoldo e Alberto Mondadori.

Fan, C., & Li, Y. (2022). *Network extraction and analysis of character relationships in Chinese literary works*. Computational Intelligence and Neuroscience, 2022(1), 7295834.

Flanagan, J. T. (1953). *The Spoon River Poet*. Southwest Review, 38(3), 226-237.

Flanagan, J. T. (1974). *Edgar Lee Masters: the Spoon River poet and his critics*. *Scarecrow Press.*

Francis, D. W. (2020). *Excavating Freytag's Pyramid: Narrative, identity and the museum visitor experience* (Doctoral dissertation, UCL (University College London)).

Franzosi, R. (2010). *Quantitative narrative analysis* (No. 162). Sage.

Franzosi, R., De Fazio, G., & Vicari, S. (2012). *Ways of measuring agency: an application of quantitative narrative analysis to lynchings in Georgia (1875–1930)*. Sociological Methodology, 42(1), 1-42.

Gervás, P., Lönneker-Rodman, B., Meister, J. C., & Peinado, F. (2006, May). *Narrative models: Narratology meets artificial intelligence*. In International Conference on Language Resources and Evaluation. Satellite Workshop: Toward Computational Models of Literary Analysis (pp. 44-51). LREC: Génova.

Goldenberg, D. (2021). *Social network analysis: From graph theory to applications with python*. arXiv preprint arXiv:2102.10014.

Gregory, H., & Zaturenska, M. (1946). *A history of American poetry, 1900-1940*. Harcourt, Brace and Company.

Gullo, A. (2019). *A 'garland'of epitaphs from the American Midwest*. Studi Classici e Orientali, 65(2), 407-422.

Hallwas, J. E. (1992). *Introduction*. In E. L. Masters, *Spoon River anthology: An annotated edition*. *University of Illinois Press*. 1-79.

Hartley, L. T. (1963). *Spoon River Revisited* (No. 1). Ball State Teachers College.

Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., Melo, G. D., Gutierrez, C., ... & Zimmermann, A. (2021). *Knowledge graphs*. ACM Computing Surveys (Csur), 54(4), 1-37.

Jewett, E. (1942). *The social philosophy of Vachel Lindsay, Carl Sandburg, and Edgar Lee Masters* (Doctoral dissertation, Boston University).

Jim, J. R., Talukder, M. A. R., Malakar, P., Kabir, M. M., Nur, K., & Mridha, M. F. (2024). *Recent advancements and challenges of nlp-based sentiment analysis: A state-of-the-art review*. Natural Language Processing Journal, 100059.

John, M., Baumann, M., Schuetz, D., Koch, S., & Ertl, T. (2019, April). *A visual approach for the comparative analysis of character networks in narrative texts*. In 2019 IEEE Pacific Visualisation Symposium (PacificVis) (pp. 247-256). IEEE.

Joyeux-Prunel, B. (2024). *Digital humanities in the era of digital reproducibility: towards a fairest and post-computational framework*. International Journal of Digital Humanities, 6(1), 23-43.

Kamil, I., & Shah, N. (2024). *Comparison of lexicon-based method, machine learning and chatgpt on sentiment analysis of big cap and small cap companies in united state indexes*. Journal of Theoretical and Applied Information Technology, *102*(16).

Kheiri, K., & Karimi, H. (2023). *Sentimentgpt: Exploiting gpt for advanced sentiment analysis and its departure from current machine learning*. arXiv preprint arXiv:2307.10234.

König, J. C. (2020). *The never-ending story teller–a narratological genealogy of storytelling in marketing and management*. International Review of Management and Marketing, 10(5), 127.

Krämer, S. (2023). *Should we really 'hermeneutise'the Digital Humanities? A plea for the epistemic productivity of a 'cultural technique of flattening'in the Humanities*. Journal of Cultural Analytics, 7(4).

Krijanskaia, D. (2008). *A Non-Aristotelian Model: Time as Space and Landscape in Postmodern Theatre*. Foundations of science, 13, 337-345.

Kwon, H. C., & Shim, K. H. (2017). *An improved method of character network analysis for literary criticism: a case study of< hamlet>*. International Journal of Contents, 13(3), 43-48.

Labatut, V., & Bost, X. (2019). *Extraction and analysis of fictional character networks: A survey*. ACM Computing Surveys (CSUR), 52(5), 1-40.

Lee, O. J., & Jung, J. J. (2020). *Story embedding: Learning distributed representations of stories based on character networks*. Artificial Intelligence, 281, 103235.

Li, J., Sun, A., & Joty, S. R. (2018, July). *SegBot: A Generic Neural Text Segmentation Model with Pointer Network*. In IJCAI (pp. 4166-4172).

Lossio-Ventura, J. A., Weger, R., Lee, A. Y., Guinee, E. P., Chung, J., Atlas, L., ... & Pereira, F. (2024). *A comparison of ChatGPT and fine-tuned Open Pre-Trained Transformers (OPT) against widely used sentiment analysis tools: sentiment analysis of COVID-19 survey data*. JMIR Mental Health, 11, e50150.

Lu, Q., Sun, X., Long, Y., Gao, Z., Feng, J., & Sun, T. (2023). *Sentiment analysis: Comprehensive reviews, recent advances, and open challenges*. IEEE Transactions on Neural Networks and Learning Systems.

Mao, Y., Liu, Q., & Zhang, Y. (2024). *Sentiment analysis methods, applications, and challenges: A systematic literature review*. Journal of King Saud University-Computer and Information Sciences, 102048.

Masters, E. L. (1920). *Domesday book*. Macmillan.

Masters, E. L. (1992). *Spoon River anthology: An annotated edition* (J. E. Hallwas, Ed.). University of Illinois Press.

Masters, E.L. (1968). *The New Spoon River*. *Macmillan*.

Min, S., & Park, J. (2019). *Modeling narrative structure and dynamics with networks, sentiment analysis, and topic modeling*. PloS one, 14(12), e0226025.

Murphy, P. D. (1988). *The Dialogical Voices of Edgar Lee Masters'" Spoon River Anthology"*. Studies in the Humanities, 15(1), 13.

Narayana, C. K. (1988). *Revolt from the Grave: Spoon River Anthology by Edgar Lee Masters*. Midwest Quarterly, 29(4), 438-47.

Narveson, R. D. (1962). Edgar lee master's'spoon river anthology':background, composition, and reputation (Doctoral dissertation, The University of Chicago).

Nelis, D., Forstall, C. W., & Milić, L. G. (2017). *Intertextuality and narrative context: digital narratology?*. hal-01480773f.

NOE, M. (2021). *Still revolting after all these years: the power of metaphor in "the revolt from the village: 1920"*. being a collection of essays commemorating the 100th anniversary of, 182.

O'Halloran, K. (2024). *Digital assemblages with AI for creative interpretation of short stories*. Digital Scholarship in the Humanities, 39(2), 657-689.

Pimentel, D. (2023). *The Agency of the Other and the Question of Violence: Otherwise than Levinas*. Philosophy International Journal. 6 (1).

Piper, A. (2023, December). *Computational narrative understanding: a big picture analysis*. In Proceedings of the Big Picture Workshop (pp. 28-39).

Piper, A., So, R. J., & Bamman, D. (2021, November). *Narrative theory for computational narrative understanding*. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (pp. 298-311).

Prince, G., & Noble, A. (1991). *Narratology, narrative, and meaning*. Poetics Today, 12(3), 543-552.

Ries, T., van Dalen-Oskam, K., & Offert, F. (2024). *Reproducibility and explainability in digital humanities*. International Journal of Digital Humanities, 6(1), 1-7.

Rochat, Y. (2014). *Character networks and centrality* (Doctoral dissertation, Université de Lausanne, Faculté des sciences sociales et politiques).

Ruediger, D., & MacDougall, R. (2023). *Are the Humanities Ready for Data Sharing?.* ITHAKA S+R. http://www.jstor.org/stable/resrep49500

Russell, H. K. (2001). *Edgar Lee Masters: A Biography*. University of Illinois Press.

Sahle, P. (2015). *Digital Humanities? Gibt's doch gar nicht*. Grenzen und Möglichkeiten der Digital Humanities.

Santana, B., Campos, R., Amorim, E., Jorge, A., Silvano, P., & Nunes, S. (2023). *A survey on narrative extraction from textual data*. Artificial Intelligence Review, *56*(8), 8393-8435.

Sharma, N., Verma, J., Latha, R., & Dharanish, J. (2024). *Quantitative Analysis of Literary Texts: Computational Approaches in Digital Humanities Research*. Educational Administration: Theory and Practice, 30(5), 5234-5240.

Snijders, T. A. (2012). *Transitivity and triads*. University of Oxford. Online: http://www. stats. ox. ac. uk/snijders/Trans_Triads_ha. pdf (last access 15-Sep-2024).

Sorrentino, G. (2020). *Medium Aevum in fabula: Il romanzo e il fascino del Medioevo*. Didattica della storia–Journal of Research and Didactics of History, 2(1S), 333-352.

Sosič, A. Z. (2022). *A Theory of Narrative*. Cambridge Scholars Publishing.

Srinivasan, V., & Power, A. (2022). *Character extraction and character type identification from summarised story plots*. Journal of Computer-Assisted Linguistic Research, 6, 19-41.

Sugishita, K., & Masuda, N. (2023). *Social network analysis of manga: similarities to real-world social networks and trends over decades*. Applied Network Science, 8(1), 79.

Swenson, M. (1962). *Introduction*. In *Spoon River Anthology*. E. L. Masters. Macmillan.

Tebaldi, M., Calaresu, M., & Purpura, A. (2019). *The power of the president: a quantitative narrative analysis of the diary of an Italian head of state (2006–2013)*. Quality & Quantity, 53, 3063-3095.

Thakre, S. S. (2024*). Echoes of the future: bridging tradition and innovation in digital humanities and english literature*. *ShodhGyan-NU: Journal of Literature and Culture Studies*, *2*(1), 16-23.

Tiefenbach Keller, S. (2024). *Towards a culture of replication in the Digital Humanities: an understanding of transparency and openness practices in published DH papers* (Master's thesis, McGill University)

Todorov, T., & Weinstein, A. (1969, October). *Structural analysis of narrative*. In NOVEL: A forum on fiction (Vol. 3, No. 1, pp. 70-76). Duke university press.

Trocchianesi, R., & Bollini, L. (2023). *Design, Digital Humanities, and Information Visualisation for Cultural Heritage*. Multimodal Technologies and Interaction, 7(11), 102.

Trocchianesi, R., & Elli, T. (2023). *Design e Digital Humanities. Approcci interdisciplinari e casi studio progettuali nell'ambito dell'Information Visualisation*. *RIVISTA DI RICERCA E DIDATTICA DIGITALE*, *3*, 9-30.

Utard, J. (2024). *Navigating the Fluidity of American Poetry Collections: Editorial, Critical, and Pedagogical Issues*. Transatlantica. Revue d'études américaines. American Studies Journal, (1).

Uyar, A., & Aliyu, F. M. (2015). *Evaluating search features of Google Knowledge Graph and Bing Satori: entity types, list searches and query interfaces*. Online Information Review, 39(2), 197-213.

VanWagenen, J. (2019, November). *Masters vs. Lee Masters: The legacy of the Spoon River author between Illinois and Italy*. In *Forum Italicum* (Vol. 53, No. 3, pp. 679-698). Sage UK: London, England: SAGE Publications.

Vercellone-Smith, P., Jablokow, K., & Friedel, C. (2012). *Characterizing communication networks in a web-based classroom: Cognitive styles and linguistic behavior of self-organizing groups in online discussions*. Computers & Education, *59*(2), 222-235.

Wang, L., Ang, L. H., & Halim, H. A. (2020). *A systematic literature review of narrative analysis in recent translation studies*. Pertanika Journal of Social Sciences & Humanities, 28(1), 1-16.

Wasserman, S. (1994). *Social network analysis: Methods and applications*. The Press Syndicate of the University of Cambridge.

Windhager, F., & Mayr, E. (2024). *Digital Humanities and Distributed Cognition: From a Lack of Theory to its Visual Augmentation*. Journal of Cultural Analytics, 7(4).

Xue, S., Lüdtke, J., Sylvester, T., & Jacobs, A. M. (2019). *Reading shakespeare sonnets: combining quantitative narrative analysis and predictive modeling—an eye tracking study*. Journal of Eye Movement Research, 12(5).

Yusof, N., & Rahman, A. A. (2009, April). *Students' interactions in online asynchronous discussion forum: A Social Network Analysis*. In 2009 International Conference on Education Technology and Computer (pp. 25-29). IEEE.

ONLINE RESOURCES

Spoon River anthology: The definitive online edition. (12-Nov-2007). Retrieved from https://web.archive.org/web/20071112102800/http://spoonriveranthology.net/spoon/river/