PNAS RESEARCH ARTICLE | NETWORK AND NLP ANALYSIS

# Character Evolution in The Lord of the Rings Trilogy: A Network, Sentiment, and Linguistic Analysis

Konstantinos Lykostratis[1,2], Trinity Sara McConnachie Evans[1,3], and Celinna Arellano Cruz[1,4]

This manuscript was compiled on December 10, 2025

**Understanding how epic narratives maintain coherence across multiple intersecting storylines remains a fundamental challenge. We tackle this by applying network science, sentiment analysis, and linguistic modeling to The Lord of the Rings film trilogy, analyzing 2,260 dialogue exchanges among the 20 main characters. Our network analysis reveals a tightly connected structure (average path length 1.268, diameter 2) with disassortative mixing, meaning central characters bridge different groups rather than clustering together. Tracking network evolution shows expansion from 119 to 196 edges as the Fellowship fractures, then contraction to 187 edges as storylines reconverge. Surprisingly, sentiment analysis reveals overall neutrality rather than expected emotional extremes, with Gollum scoring as both the most volatile and most linguistically positive character due to his obsessive relationship with the Ring. Our findings show that computational methods can reveal hidden narrative patterns, quantify character relationships, and trace emotional arcs in ways traditional analysis struggles to capture at scale.**

## Significance

This study investigates how the main characters of The Lord of the Rings film trilogy evolve through their interactions, emotional trajectories, and linguistic patterns. Using tools such as social network analysis, sentiment analysis, and vocabulary modeling, we quantify changes in relationships, emotional tone, and language across the three films. Our central question asks how these interaction patterns, emotional shifts, and linguistic signatures reveal broader narrative structure, character roles, and underlying themes. Although applied here to a fictional universe, the analytical methods are widely used in scientific fields to study complex systems. This work illustrates how computational approaches can deepen our understanding of human narratives and cultural dynamics.

The Lord of the Rings trilogy presents one of the most complex narrative structures in modern cinema, with an ensemble cast navigating diverging and converging plotlines across three films. Understanding how such narratives maintain coherence while managing multiple character arcs remains a fundamental challenge. Traditional literary approaches rely on close reading, but struggle to capture the full complexity of large-scale interaction patterns across extended narratives.

Network science offers a powerful framework for analyzing complex relational systems. As Barabási demonstrates in Network Science(1), real-world networks exhibit universal organizational principles. Recent work has applied these principles to literary texts, treating characters as nodes and their interactions as edges. However, most studies focus solely on network structure, neglecting the emotional and linguistic dimensions that give narratives their depth.

We address this gap by combining network analysis with sentiment analysis and natural language processing to examine character evolution in the Lord of the Rings trilogy. Through analysis of 2,260 dialogue exchanges, we investigate how interaction patterns, emotional trajectories, and vocabulary use reveal the underlying structure of Tolkien's narrative. Our findings demonstrate that network centrality identifies narratively important characters more reliably than dialogue counts, that emotional trajectories follow systematic patterns tied to plot structure, and that characters possess distinctive linguistic signatures reflecting their narrative roles.

## Results

### Network Analysis.

***Network Construction and Structure.*** We constructed our network by treating the main 20 characters (identified by dialogue count) as nodes and their consecutive dialogue exchanges as directed edges. Using consecutive quote pairings from 2,260 total dialogue exchanges, we built interaction networks with NetworkX, weighting edges by interaction frequency. It's important to note that this

Author affiliations: [a]Technical University of Denmark (DTU); [b]Technical University of Denmark (DTU); [c]Technical University of Denmark (DTU)

Trinity: Results-Sentiment Analysis, Results-Language Pattern Analysis, Discussion-Limitations, Methods. Coded initial Network construction and network analysis.
Celinna: Abstract, Introduction, Results-Network Analysis. Coded additional network analysis and initial sentiment analysis.
Konstantinos: Significance Statement, Discussion. Coded advanced sentiment and network analysis (temporal evolution, centrality, custom sentiment dictionary), wordcloud visualization, and responsible for combining code contributions into one notebook, adding explainer details

[1]Authors contributed equal portions to the linked code notebook

[2]E-mail: s250269dtu.dk
[3] E-mail: s252753dtu.dk
[4] E-mail: s252660dtu.dk

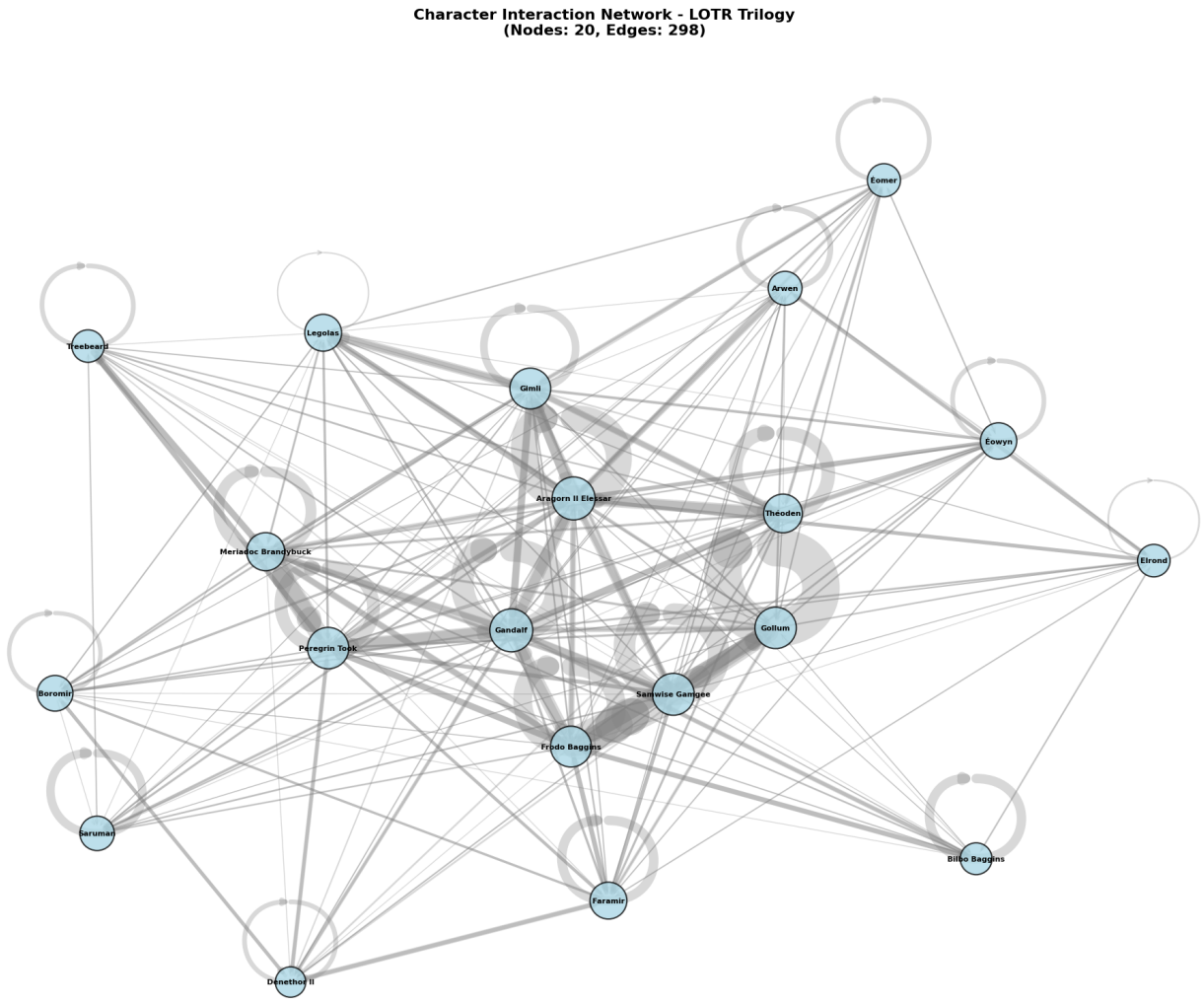


**Fig. 1.** Character interaction network across the LOTR trilogy. Nodes represent the 20 main characters, with edges indicating directed interactions between them. Node size reflects connection prominence, while edge thickness represents interaction frequency. Self-loops indicate internal dialogues or monologues. The network's high connectivity (298 edges) demonstrates the narrative's interwoven character relationships, with central characters like Gandalf, Aragorn, and Frodo positioned at key bridging locations.

method considered any two consecutive lines spoken by main characters as an interaction, meaning 3'character interactions may not have been fully captured.

The resulting network contained 298 directed edges in a single giant connected component (Figure 1). The average degree was 14.90 with a maximum of 19, proving that main characters maintained broad interaction patterns rather than clustering into separate subgroups. The average shortest path length was 1.268 and a diameter of 2, meaning that on average, characters are just slightly more than 1 step apart and that the longest shortest path in the network is 2 steps apart. The degree assortativity coefficient of –0.168 showed disassortative mixing, meaning that hub characters preferentially interacted with less connected characters rather than other hubs, creating a structure where central protagonists bridge different character groups.

***Character Centrality.*** Betweenness centrality measured how often each character served as a bridge between other characters. As visible in the network structure (Figure 1), the highest scores belonged to Gandalf (0.0388), Aragorn (0.0388), Sam (0.0329), Gollum (0.0329), and Pippin (0.0268), revealing their roles as key bridges between narrative threads. These characters occupy central positions in the network visualization, with numerous connections radiating to other nodes.

Degree and betweenness centrality showed a strong correlation ($r = 0.823$), meaning characters with many connections typically also bridged different network regions. The imperfect correlation indicated role differentiation: some high-degree characters operated within dense clusters, while others with fewer connections occupied strategic bridging positions.

***Network Evolution.*** By building separate networks for each film, we discovered how character interactions shifted across the trilogy. Fellowship of the Ring had the smallest network with 14 characters and 119 edges, reflecting the focused narrative as the core fellowship formed and traveled together.

The network expanded in the Two Towers to 19 characters and 196 edges, making it the largest configuration of the trilogy. This expansion captures both the introduction

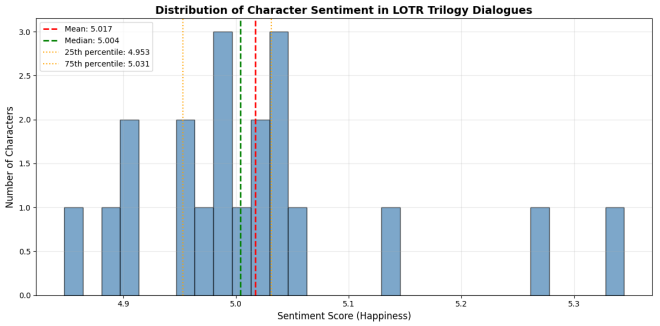


**Fig. 2.** Distribution of character sentiment in the LOTR trilogy.

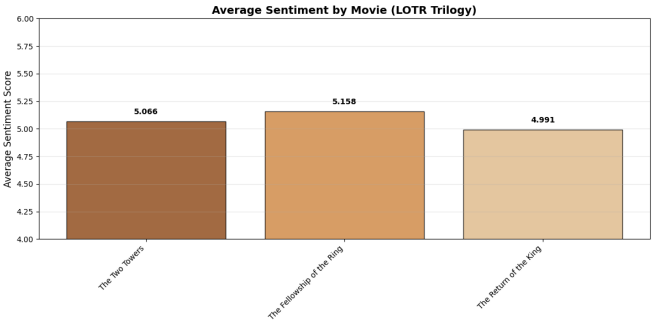


**Fig. 3.** Average sentiment by movie in the LOTR trilogy.

of new major characters (Éomer, Théoden and Faramir) and the fellowship's fracturing into parallel storylines. The increased edge count reflected multiple groups having separate conversations simultaneously.

Return of the King kept the same 19 characters but contracted to 187 edges. Despite the constant cast size, the reduction in total interactions indicated the parallel storylines beginning to reconverge as the narrative moved toward its unified conclusion.

**Sentiment Analysis.** To calculate the sentiment for each character, we created a joined list of each character's dialogues from the trilogy and ran tokenization and sentiment functions on each. We then analyzed the emotional range of each character by performing a volatility analysis, and performed analyses on their evolution over time.

***Initial Analysis.*** After performing our sentiment analysis of the characters and movies we were surprised to find that the sentiment of the LOTR trilogy remained largely neutral, with a mean sentiment of the characters being 5.017 on a scale from 0 to 10 (See Figure 2). Sentiments were approximately normally distributed around 5.0, suggesting a balanced emotional tone rather than an overtly positive or negative one. This could reflect a more even tone in the movies, between despair and hope. The main characters had a slightly positive tone, but were no exception to the overall neutrality. Interestingly, the character with the strongest sentiment score was Gollum, at 5.344, also making him technically the most positive character. This may seem counterintuitive given that he is a movie villain and antagonist, but those familiar with his story will recognize that it is his corrupting obsession with the ring, which he speaks of positively, which shapes his character's arc. This suggests that our analysis captures his linguistic positivity when referring to 'my precious,' despite the dark psychological undertones.

***Emotional Volatility.*** To measure emotional consistency we calculated the standard deviation from the mean sentiment score for each character. The most emotionally volatile character in the trilogy was Gollum, whose standard deviation from his mean sentiment was 1.128. A contrasting main character such as Frodo Baggins was more emotionally consistent, at 0.628. This aligns with each of their characters journeys throughout the movies; Gollum's continued descent into obsession and unstable nature caused by the ring, and Baggins consistent bravery and commitment to his friends. Emotional volatility seemed unaffected by character narrative importance in LOTR in

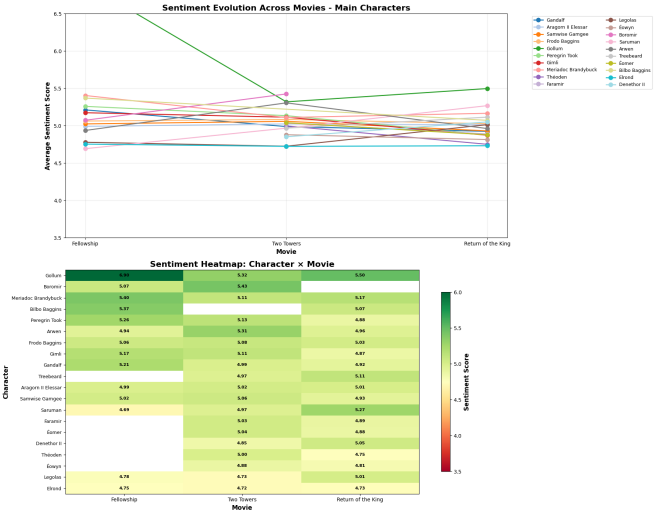


**Fig. 4.** Sentiment evolution of characters in LOTR.

general; the more minor characters and the protaganists did not vary greatly in their volatility. For example, Eowyn and Elrond, two characters who remain largely on the outskirts of the main story, had emotional volatility of 1.01 and 0.645 respectively.

***Sentiment Evolution.*** As the characters storylines continued through the trilogy, the sentiment scores of their dialogue also changed. As can be seen in Figure 3 the average sentiment in each of the movies varied minimally. The character with the greatest positive sentiment change from the first movie to the last was the character Saruman, who had an average sentiment score of 4.69 in the first movie, and 5.27 in the last (See Figure 4). This increase in positive linguistic patterns is interesting given his characters fall from good to evil *(spoilers)*. The character with the greatest negative sentiment change was Gollum, whose sentiment score dropped by 1.40 from the first movie to the last. This is not surprising given his descent to madness.

**Language Pattern Analysis.**

***WordCloud.*** Sentiment analysis was informative on the linguistic sentiment of the dialogue throughout the movies, but knowing that the sentiment wasn't quite consistent with the characters meaning and tone, we wanted to see what language was actually being used. To better visualize and understand language usage in the LOTR trilogy, we built a variety of WordClouds of the happiest

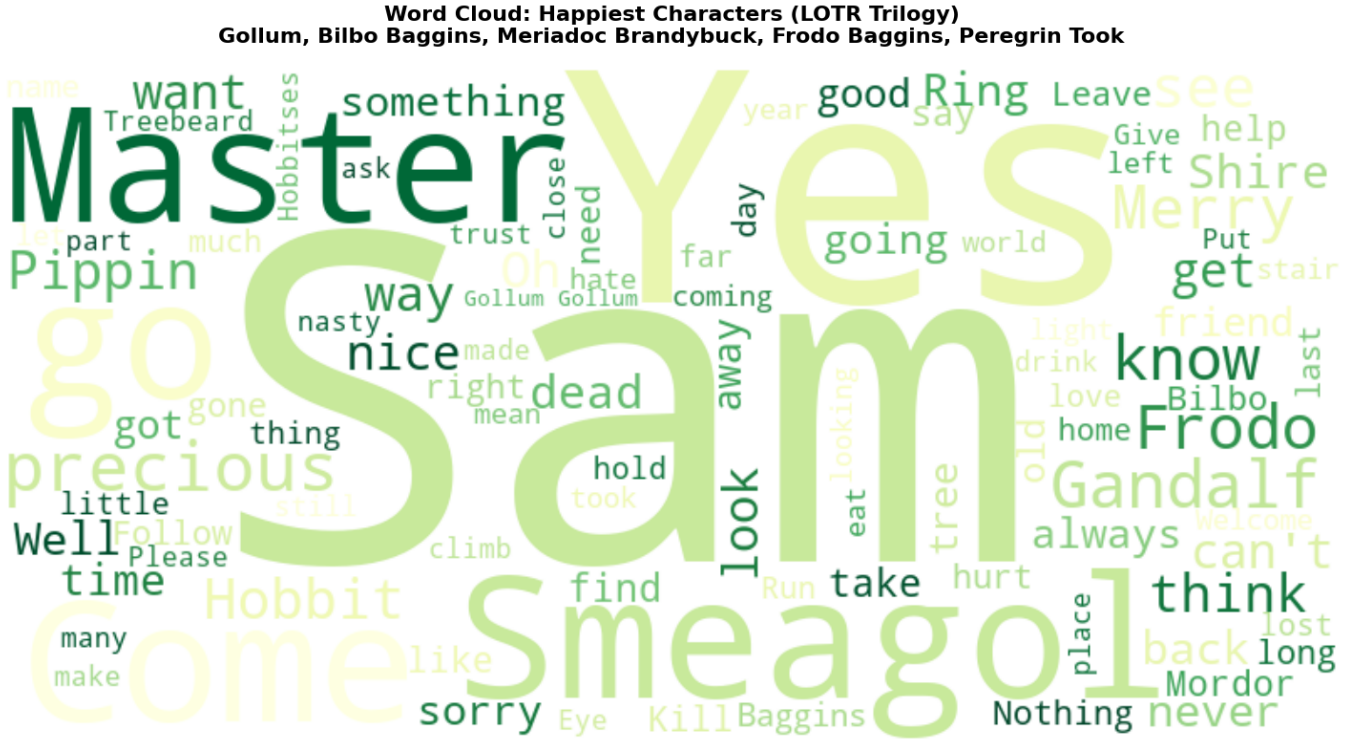


**Fig. 5.** Wordcloud representing the language use of the 5 happiest characters in the LOTR trilogy.

and saddest characters. As can be observed in Figure 5, the WordCloud for the 5 "happiest" characters (most positively ranked sentiment-wise) included a variety of positive language but featured words used by Gollum quite heavily (such as "Master" and "Precious"). This confirmed to us that the positive ranking for Gollum was more connected to the technical linguistic positivity of his dialogue than any actual goodness in his character.

## Discussion

By combining network analysis, sentiment analysis, and linguistic modeling, this study provides a multidimensional look at how the main characters of *The Lord of the Rings* evolve across the trilogy. The interaction network shows a clear shift from the tightly connected Fellowship in the first film to a set of smaller groups following different narrative paths in the later films. This mirrors the story's growing complexity and highlights how characters like Gandalf and Aragorn continue to play connecting roles, while Frodo gradually becomes more isolated as the Ring's burden intensifies.

The emotional patterns found through sentiment analysis add depth to this picture. Frodo's increasingly negative tone, Sam's steady optimism, and Gollum's emotional volatility all align closely with their narrative journeys. Linguistic analysis further distinguishes characters by the themes and ideas reflected in their speech, from Gandalf's language of guidance to Gollum's fixation on the Ring. These quantitative perspectives support and enrich well-known qualitative interpretations of the trilogy.

This study did not come without challenges and limitations. When building our network edges we followed similar approaches seen in character network analysis (2) and focused on dialogue-based interactions between characters. Our method considered any two consecutive lines spoken by main characters as an interaction; the limitations of this method meant that 3+ character interactions may not have been properly accounted for, and some consecutive lines may have been mistakenly categorized as interactions. For our sentiment analysis portion, we built a 113 word fall back sentiment dictionary in the case that the LabMT list could not be retrieved. This list was limited in scope and as such it is worth considering the potential impact on our results.

These findings illustrate how computational tools can reveal structure and meaning within a story that might not be immediately visible through traditional analysis alone. Although our approach is limited by its reliance on dialogue and simplified sentiment dictionaries, it demonstrates the value of combining multiple analytical perspectives. Future work could incorporate more detailed scene information, advanced language models, or temporal network analysis to capture character development even more fully and to compare different adaptations or related works.

## Materials and Methods

**0.1. Dataset and Processing.** For our project we used data from the Lord of the Rings API **"The One"**(3), which contains character and dialogue information from all of The Lord of the Rings (LOTR) movies. In order to narrow down our scope, we processed the data of only the original LOTR trilogy: *The Two Towers*, *The Fellowship of The Ring*, and *The Return of The King*. After cleaning, our dataset contained 933 characters and 2260 quotes of movie dialogue.

**0.2. Network Construction and Visualization.** We constructed directed character networks from 2,260 dialogue exchanges, creating edges between consecutive speakers (2). Edge weights represented interaction frequency. Network metrics were computed using NetworkX.

**0.3. Sentiment Analysis.** Character sentiment was quantified using a custom 113-word dictionary with happiness scores ranging from 0-10. Detailed methodology is available in the supplementary notebook found at this link: GitHub Link.


1. AL Barabási, *Network Science*. (Cambridge University Press), (2016).
2. V Labatut, X Bost, Extraction and analysis of fictional character networks: A survey. *ACM Comput. Surv.* **52** (2019).
3. The One API, The one api: The Lord of the Rings api (https://the-one-api.dev) (2024) Accessed: 2025-12-10.