

# Textual Analysis of Topic Titles for a Game Character

LIU Taijiang 3036025836

## 1. Introduction

Genshin Impact is an action role-playing game developed and published by miHoYo. The game takes place in the fictional world of Teyvat, where players assume the role of a traveler searching for their lost sibling. Since its launch, Genshin Impact had been hugely popular in the global gaming market due to its open-world gameplay system and character-driven storytelling. Players' perception of the characters is crucial to the game's success as it serves as the primary source of income and revenue. Excellent character story and high levels of discussion intensity can lead to higher profits and recharge rates. Before the release of a new playable character, players will discuss on various game forums, video-sharing platforms, and other online communities. In the weeks leading up to the character's release, there may even be insider leaks providing information on the character's abilities and skills. Character skills are a crucial element of the game, with discussions on popular topics such as skill strength, elemental attributes, character roles, and team composition. The NGA (National Geographic of Azeroth) forum, a popular gaming forum in China, features a dedicated Genshin Impact section that hosts extensive discussions by a large number of players.

Players' attention and positive or negative evaluations before and after character release are likely to differ for differences between leaked information and actual character details. Even as the game continues to update, the sentiment of its reviews (positive or negative) can continue to change and affect its future sales performance. This paper aims to explore the differences in player evaluations caused by information discrepancies before and after character release, as well as changes in the sentiment of comments during the entire product cycle. Ye Lan is a character that was released on May 31, 2022. Due to her powerful skill design and controversial image design, her popularity has remained a persistent topic of discussion. According to Baidu Index, her popularity remained high in the month before her release, reached its peak in the first two weeks after her release, and reached a second peak during her first return sale on February 7, 2023. So, I chosen Ye Lan as the subject of this article and investigated the differences in topic content and sentiment one month before and after her release, as well as the differences in topic content and sentiment before and after the release of the subsequent version of XUMI.



Figure1. Baidu Index of Ye Lan(夜兰)

## 2. Methodology

### 2.1 Data resource and processing

The data was obtained from the "Genshin Impact" sub-forum on the NGA forum. Due to the difficulty in processing comment data, only the titles and release dates related to "Ye Lan" were collected with Python. The Python code is included in the attachment. The preprocessed data includes only two variables: "title"- the text data containing titles related to Ye Lan, and "date" - the corresponding date of the titles.

### 2.2 Word cloud

To conduct a preliminary analysis of the content related to Ye Lan, I created word clouds for the following four time periods: (a) one month before the character was released (April 4, 2022 to May 31, 2022), (b) one month after the character was released (May 31, 2022 to June 30, 2022), (c) before the character entered the XUMI version (April 4, 2022 to August 24, 2022), and (d) after the character entered the XUMI version (August 24, 2022 to April 5, 2023). When performing tokenization, I added a self-made dictionary containing relevant terms such as the character's name, skills, game props, and maps to jiebaR, which is included in the attachment.

### 2.3 LDA-tuning, Cross Validation and LDA

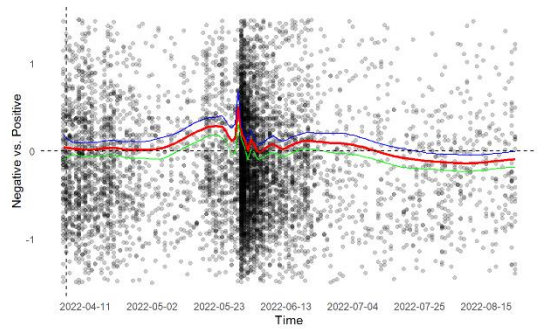
To further explore the comments about Ye Lan throughout the complete product cycle, I chose to use LDA for topic analysis. LDA-tuning and 5-fold Cross Validation were used to determine the number of topics - K value.

### 2.4 Sentiment analysis

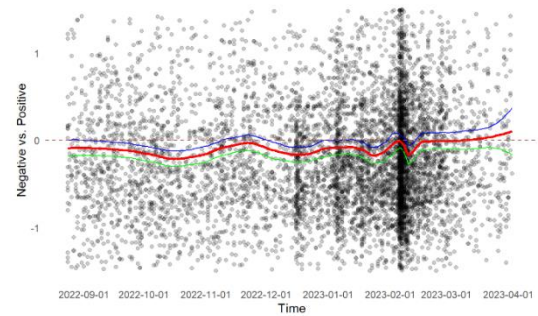
To further explore the overall sentiment of comments on Ye Lan, the author conducted Latent Semantic Scaling sentiment analysis on the following four time periods: (a) one month before the character was released (April 4, 2022 to May 31, 2022), (b) one month after the character was released (May 31, 2022 to June 30, 2022), (c) before the character entered the XUMI version (April 4, 2022 to August 24, 2022), and (d) after the character entered the XUMI version (August 24, 2022 to April 5, 2023). The NTUSD (National Taiwan University Sentiment Dictionary) is a Chinese sentiment lexicon that contains words and phrases with positive or negative







**Figure12. Sentiment analysis before XUMI**



**Figure13. Sentiment analysis after XUMI**

The graph shows that the overall sentiment towards the character Yan Lan in Genshin Impact shifted from positive to negative and back to positive again. Before the character's release, players were generally optimistic about Yan Lan and the sentiment continued to rise in the month leading up to the release, reaching its peak in the week before. However, after the character was officially released, the sentiment dropped to a neutral attitude and remained relatively weakly positive in the following month. Considering the scale of Genshin Impact and huge player base, maintaining a generally neutral evaluation towards a character can be a sign of successful product shaping. "XUMI", is an innovative version of Genshin Impact that introduced a completely new system, which had an impact on the existing system. The graph also shows that Yan Lan's comment sentiment had a negative value in the month before the release of XUMI, and it gradually turned positive after the first return sale on February 7th, 2023.

Analyzing discussion topics related to a game product can provide information for product and sales decisions. However, the trend in sentiment analysis discussed in this article has various bias, such as intentional leak of character news and hype by the game company, asymmetry between the discussion groups in game forums and the actual player groups, and the nature of data that makes it difficult to obtain deeper information.