

**Group Project Idea: Group F**

Yaran Lu(S5482488), Mengqi Li(S5479606),  
Shuwei Wang(S5466121), Zixuan Wu(S5483662),  
Yuxin Huang(S5574242), Jiaming Chen(S5692302)

Faculty of arts, University of Groningen

LHU011M05: Digital Humanities: Tools and Methods

Dr. Federico Piazola

Dr. Lucas van der Deijl

December 18, 202

## **Project Introduction and Goal**

This project aims to employ Digital Humanities tools for literary analysis, specifically focusing on "THE DREAM OF THE RED CHAMBER," a globally renowned Chinese classical novel. Through applying digital tools, we aim to unveil new dimensions of this literary masterpiece. Lin Tai-yü, based on prior reading experiences, emerges as a character of exceptional uniqueness within "THE DREAM OF THE RED CHAMBER." This project endeavors to concentrate on Lin Tai-yü's discourse, utilizing quantitative data to delve into her distinct linguistic artistry and personality traits. Visual representations derived from this analysis will offer an insightful exploration of her unique linguistic expressions and individual characteristics.

## **Research Question**

How do the language styles differ between Lin Tai-yü and other characters in the dialogue text of the Chinese classic "THE DREAM OF THE RED CHAMBER"? Exploring the specific manifestations of these distinctions and analyzing whether these distinct language features accurately portray Lin Tai-yü's character will be our research question throughout the project.

## **Dataset Description**

Due to "THE DREAM OF THE RED CHAMBER" will be the object of study, and Lin Tai-yü' is one of the main characters of this Chinese masterpiece. Therefore, the English version of its original article is our metadata stored in text format. With the purpose of comparing Lin Tai-yü to other characters, our group will apply Python to process the metadata and abstract Lin Tai-yü's discourse only. In conclusion, two primary datasets are used in this project: the translated "THE DREAM OF THE RED CHAMBER" and the abstraction of Lin Tai-yü.

## **Tools Used and Methods**

In order to accomplish our research question, which emphasizes the analysis of words and characters, Voyant and Stylo will be two dominant tools applied within the project. While the former tool helps list different term or collocation usage habits, the latter provides various intuitive diagrams of stylistic characteristics, including word frequencies, sentence lengths, and linguistic markers. The output of this project is presented in PDF, and the contents are comprised of the scripts of the diagram processing datasets and compatible text annotations.

### **Project Limitations**

Moreover, our group does realize the limitation of our project while applying Python code to abstract Lin Tai-yü. We are unable to ensure the abstraction is accurate and complete due to the uncertainty in describing this character in literature. However, although our group is not capable of solving this problem for the time being, we are determined to minimize the possibility of potential risks and safeguard the maximum accuracy of results.