The project proposal

1. Name of your project and a short description.

Project title: **"Analysis of Chinese and English Language Perspectives on the Palestinian-Israeli Conflict"**

Introduction: This project aims to analyze and compare reports and perspectives on the Palestinian-Israeli conflict in Chinese and English media. By capturing and analyzing data from Chinese and English reports on the Palestinian-Israeli conflict in Wikipedia and Google News, the aim is to reveal reporting biases and differences in opinions under different language backgrounds.

2. What problem are you trying to solve, which question are you trying to answer?

The question addressed is: Are there obvious differences in views or biases between Chinese and English media reports on the Palestinian-Israeli conflict?

The question I want to answer is: In what ways do these differences or biases manifest themselves? How can these differences be revealed through data analysis?

3. How do you intend to collect the data and where is it coming from?

Data Sources: 1. Chinese and English versions of Wikipedia's entry on the Palestinian-Israeli conflict. 2. Google News reports on the Israeli-Palestinian conflict in Chinese and English.

Data crawling method: Use the `scrape_wikipedia_page` and `save_news_to_csv` functions to scrape relevant data from Wikipedia and Google News RSS feeds.

4. What analysis will you do and what kind of visualizations are you going to create?

Data preprocessing: including word segmentation, removal of stop words, etc.

Analysis type:

1. Statistical analysis of word frequency and keywords: identify the most frequently occurring words and keywords.

2. Time series analysis: observe the changing trend of reports over time.

3. Text similarity analysis: Compare the similarity of the content of Chinese and

English reports.

4. Sentiment analysis: Analyze the emotional tendencies in reports.

5. Topic modeling analysis: Identify the main topics in the report.

6. Language style analysis: Compare the style characteristics of reports in different languages.

Visualization: Use charts and graphs to display analysis results, such as bar charts, time series charts, word clouds, etc.

Summary of conclusion: Based on the analysis results, summarize the differences or biases in views between Chinese and English reports on the Palestinian-Israeli conflict.