

Multilingual history visualization

The big data problem raised in this article is to complete a multilingual history visualization website. To characterize history, we need to build a three-dimensional physical world plus time dimension in the program. We will use this space-time relationship to visualize history. We will build a graph database to store historical knowledge. Eventually we will visualize history through timelines, maps, and relationship maps.

Big data property:

Elements of history include historical figures, historical events, historical time, historical place, and historian's point of view. And we will provide historical time and place labels for each historical person, place and place.

Data sources:

1. Wikipedia history entry
2. Baidu Encyclopedia history entry
3. CHINA BIOGRAPHICAL DATABASE PROJECT (CBDB)
<https://projects.iq.harvard.edu/chinesecbdb/cbdb-api>
4. Codes of administrative regions in China <http://www.mca.gov.cn/article/sj/xzqh/>
5. Online query of the Atlas of Ancient Chinese History
<http://home.olemiss.edu/~gg/hstrymap/lishidit.htm>
6. Entry resource <https://github.com/shgongbo/allHistory>
7. Google Maps API <https://cloud.google.com/maps-platform/?hl=en>
8. Allhistory <https://www.allhistory.com/>

Database selection:

Neo4J: Graph database, suitable for current knowledge graph tasks.

MangoDB: MongoDB is a database based on distributed file storage. Easy to store text data.

Work Flow:

1. Crawl the original text and picture data. The first stage is mainly based on encyclopedia data, and the second stage introduces historical maps and figures.
2. Build database. Store text into MangoDB. Save text URLs and pictures into Neo4J.
3. Entity annotation of historical figures, events, times, places and opinions in the text through named entity recognition technology.
4. Construct historical element triples with relation extraction technology.
5. Construct historical events through complex multivariate event extraction.
6. We will generate time and geographic location labels for each historical element, and display them on the map by means of map visualization.
7. Complete the reasoning of any two historical elements based on the knowledge graph technology.