Introduction

- Multilingual knowledge bases are important for the globalization of knowledge sharing.
- However, current Wikipedia-based multilingual knowledge bases still suffer the following problems:
  - the scarcity of non-English knowledge
  - the noise in the semantic relations
  - the limited coverage of equivalent cross-lingual entities
- On the other hand, there are more and more similar large-scale non-English online wikis such as Baidu Baike and Hudong Baike.

In this demo, we present a large-scale bilingual knowledge graph named XLore, which has adequately solved the above problems.

Overview of XLore Approach

- **Data Preprocessing**: collect and clean the data sets from English Wikipedia, Chinese Wikipedia, Baidu Baike and Hudong Baike.
- **Knowledge Graph Building**: semantify the online wikis, conduct cross-lingual knowledge links, and extract the structured knowledge.
- **Knowledge Query**: construct an online system for knowledge acquisition.

Approach

- **Data Preprocessing**: collect and clean the data sets from English Wikipedia, Chinese Wikipedia, Baidu Baike and Hudong Baike.
- **Knowledge Graph Building**: semantify online wikis, conduct cross-lingual knowledge linking, and extract structured knowledge.
- **Knowledge Query**: construct an online system for knowledge acquisition.

System

- XLore harvests **856,146 classes**, **71,596 properties** and **7,854,301 instances** across English and Chinese.
- XLore supports the **keyword-based or SPARQL queries**, gives the statistical information, offers visualization demonstrations, etc.

Conclusion

- XLore is the first large-scale cross-lingual knowledge graph with balanced amount of Chinese-English knowledge.
- XLore gives a new way for building such a knowledge graph across any two languages.