Preliminary evaluation of passage retrieval in biomedical multilingual question answering

Mariana Neves[1], Konrad Herbst[1,2]

[1] Hasso-Plattner-Institute at the University of Potsdam, Germany
[2] University of Heidelberg, Germany

Introduction – Multilingual Question Answering

Multilingual question answering (QA) systems offer a variety of new possibilities to the user:
• By exploring resources in other languages;
• By allowing non-native English speakers to pose questions and receive answers in their native language.

In this work we evaluated passage retrieval in QA for three languages (English, German and Spanish) and created two parallel corpora for the evaluation.

English/German and English/Spanish corpora of questions

Based on the CLEF-ER resources:
• Medine documents (titles): 1,593,546 (English), 719,232 (German) and 247,655 (Spanish)
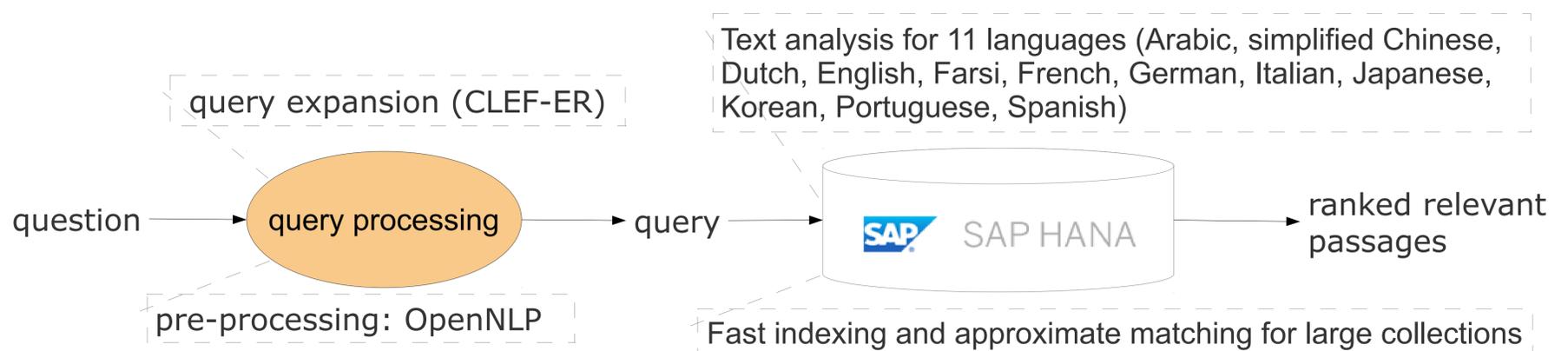• Terminology (synonyms): 1,771,498 (English), 118,902 (German), 622,638 (Spanish)

Multilingual parallel corpora for passage retrieval:
• 50 questions (English/German) and 50 questions (English/Spanish)
• Re-phrase of the original sentences using synonyms
• Available at: **https://sites.google.com/site/marianalaraneves/resources/questions.zip**

English/German (d6357751)	Which methods can be used to determine the living cell count of cariogenic microorganisms?
	Welche Methoden bieten sich zur Bestimmung der Lebendzellzahl von kariogenen Mikroorganismen an?
English/Spanish (d16888692)	What are possible drug targets for eye related infections?
	Cuáles son los posibles objetivos farmacológicos en infecciones relacionadas con el ojo?

Experiments and evaluation: passage retrieval

▪ 25 questions for test-dev and 25 questions for test
▪ Evaluation: retrieval of the original sentences from which the questions were derived

query expansion (CLEF-ER)

Text analysis for 11 languages (Arabic, simplified Chinese, Dutch, English, Farsi, French, German, Italian, Japanese, Korean, Portuguese, Spanish)

question → query processing → query → SAP SAP HANA → ranked relevant passages

pre-processing: OpenNLP

Fast indexing and approximate matching for large collections

Settings	English/German		English/Spanish	
	EN	DE	EN	ES
exact	0.04 (1)	0.04 (1)	0.01 (1)	0.09 (4)
fuzzy	0.05 (3)	0.02 (1)	0.10 (5)	0.10 (5)
+ synonyms	0.09 (3)	0.06 (2)	0.11 (4)	0.09 (6)
+ linguistic	0.10 (4)	0.04 (2)	0.03 (2)	0.08 (6)
Test	0.05 (5)	0.05 (4)	0.05 (3)	0.06 (4)

R-Precision and number of sentences correctly retrieved (in parenthesis)

▪ Passage retrieval is not a straightforward task for none of the three languages.
▪ The corpus is adequate for the task even when evaluated only on document titles.
▪ Lack of multilingual resources hindered better performance of the query expansion component.

Contact

Mariana Neves; e-mail: marianalaraneves@gmail.com

HPI Hasso Plattner Institut
IT Systems Engineering | Universität Potsdam