SQX-Lib: Developing a Semantic Query Expansion System in a Media Group

María G. Buey, Ángel Luis Garrido, Sandra Escudero, Raquel Trillo, Sergio Ilarri, Eduardo Mena
IIS Department, University of Zaragoza (Spain)

Problems in relational database searches

- **Term mismatch problem**: retrieving data where the user keywords exactly match with the terms of the documents. However, users often do not use the same words present in these documents. Therefore, it is needed performing a semantic analysis of the keywords to increase the effectiveness of the retrieval.
- **Low recall**: losing relevant information that the user is trying to get.
- **Low precision**: needless data that are not interesting for the user.

Solution & Contributions

- **Automatic Query Expansion (AQE)**: to expand the original query with other words that best capture the actual user intent or that simply produce a more useful query that is more likely to retrieval relevant documents.
- **SQX-Lib focuses on automatically and semantically expanding** the scope of the searches and fine-tune them by taking advantage of the Named Entities (NE) included in these queries.
- **Exploit linguistic and semantic techniques** to improve the quality of the keyword-based search process on relational repositories of the Heraldo Media Group.

References


Acknowledgements

This work was supported by the CICYT project TIN 2010-21387-C02-02, DGA-FSE and HERALDO DE ARAGÓN S.A.

Contact

General information: María Granados Buey (mgbuey@unizar.es)
Research group SID (http://sid.cps.unizar.es)
The GENIE Project (http://sid.cps.unizar.es/SEMANTICWEB/GENIE)