GYANI: An Indexing Infrastructure for Knowledge-Centric Tasks

Dhruv Gupta and Klaus Berberich
Max Planck Institute for Informatics
Saarland Informatics Campus, Germany

August 22, 2018

Contents. The testbed contains the set of queries compiled for our publication at CIKM 2018 titled "GYANI: An Indexing Infrastructure for Knowledge-Centric Tasks". The testbed was compiled using Yago [3], Patty [2], and the New York Times "On this Day" event repository [1]. The testbed comprises of the following files:

1. *task-fs-news.tsv*

2. *task-fs-wiki.tsv*

3. *task-ie.tsv*

4. *task-qa-news.tsv*

5. *task-qa-wiki.tsv*

6. *task-re-news.tsv*

7. *task-re-wiki.tsv*

8. *task-sq.txt*

Each file contains queries corresponding to one of the five knowledge-centric tasks used for evaluation in our publication:

- Information Extraction (IE)

- Relationship Extraction (RE)

- Question Answering (QA)

- Fact Spotting (FS)

- Semantic Search (SQ)

Citation. Kindly, cite our work in case you find it useful for your research:

Gupta D. and Berberich K.
GYANI*: An Indexing Infrastructure for Knowledge-Centric Tasks*
CIKM 2018.

References

[1] The new york times: On this day. (`https://learning.blogs.nytimes.com/on-this-day/`).

[2] Ndapandula Nakashole, Gerhard Weikum, and Fabian M. Suchanek. PATTY: A taxonomy of relational patterns with semantic types. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, EMNLP-CoNLL 2012, July 12-14, 2012, Jeju Island, Korea*, pages 1135–1145, 2012.

[3] Fabian M. Suchanek, Gjergji Kasneci, and Gerhard Weikum. Yago: A large ontology from wikipedia and wordnet. *Web Semant.*, 6(3):203–217, September 2008.