

Review of: "RelTopic: A Graph-Based Semantic Relatedness Measure in Topic Ontologies and Its Applicability for Topic Labeling of Old Press Articles"

Silvio Peroni¹

¹ University of Bologna

Potential competing interests: The author(s) declared that no potential competing interests exist.

Metadata

Title: RelTopic: A Graph-Based Semantic Relatedness Measure in Topic Ontologies and Its Applicability for Topic Labeling of Old Press Articles

Author: Mirna El Ghosh, Nicolas Delestre, Jean-Philippe Kotowicz, Cecilia Zanni-Merk, Habib Abdulrab

Submitted to: Semantic Web (resubmission after major revision request)

Preprint: <http://www.semantic-web-journal.net/content/rektopic-graph-based-semantic-relatedness-measure-topic-ontologies-and-its-applicability-0>

Review

I want to thank the authors for having addressed the main part of my comments^[1]. They have extended several parts of their work both in the text and in the experiments done. As a result, the article has been improved a lot. However, some issues (listed below) should be addressed to have the article accepted in the Semantic Web journal.

Inter-rater agreement

By reading Section 8.2, it is unclear how the authors have computed the inter-rater agreement among the three annotators and the results obtained. In particular, it is not clear how the various percentages that seem to refer to distinct aspects (i.e. 46%, 26% and 15.5%) provide an overall percentage of 82% - I cannot see how such a small percentage for each kind of annotations compose such a vast overall percentage. I perceive the same issue also for the comparison between RelTopic and the human annotators. Which particular approach has it been used to compute such percentages? Did the authors use some well-known statistics? If only simple percent agreement calculation has been used, why has this been preferred against more robust statistics such as Cohen's kappa and Fleiss' kappa, proposed for these purposes?

Full availability of data

All the data collected in the experiments, particularly those related to the quantitative evaluation, should be available online (e.g. in Zenodo) with an appropriate license. The article introduces only excerpts of these (in Table 4 and Table 5) but **all** the annotations done by the annotators should be freely available for the sake of experimental checking and reproducibility. This also applies to the outcomes illustrated in Section 8.3, where the authors said that RelTopic labelled 70% of the articles correctly. How has that percentage been computed? Where are the data confirming that 70%?

Even if it may not be mandatory yet, the Semantic Web journal (see <http://www.semantic-web-journal.net/blog/open-science-data-impending-changes-review-process-semantic-web-journal>) strongly support that, when possible (and, to me, this work is one of such cases), experimental data should be available and uploaded in persistent repositories. In addition, these data should be cited in the article.

Software availability

According to the Semantic Web policy guidelines above, to check the correctness of your implementation and foster the reproducibility of experiments, the software the authors developed (since they clarified it is open source in their response letter) should be available in a persistent repository. It is not adequate to say that it will be in the future. Also, the software is necessary to assess the quality of the work and enable the replication of the study. Thus, I would commend publishing it in a repository that makes it available to all without having to comply with specific requirements (e.g., having an account of a particular platform to access it). As a suggestion, GitHub can be used for archiving the software, and Zenodo should be adopted to have a DOI for the software (see <https://guides.github.com/activities/citable-code/>) and, consequently, to cite it in the article.

Typos

Section 2: "coming from WP2 of the project" -> "coming from WP2 of the ASTURIAS project, as shown in Fig. 2"

Section 2: "Given a corpus of ... most relevant topics from T" -> "given a corpus of articles A, a set of named entities N (represented by a set of URIs) that are collected from A (WP2), and a topical structure T, we want to find the most relevant topics described in T"

Section 2: "takes as input N the .. on N" -> "it takes as input N, i.e. the set of disambiguated named entities, and constructs T, i.e. a convenient topical structure based on N"

Section 8.2.1: "For this purpose, A... which is previously used" -> "For this purpose, we considered A as the corpus of 48 articles from Le Matin that we have introduced in Section 7"

Table 4 and 5: subscripts should be used to referring to the articles instead of A_1, etc.

References

1. ^Silvio Peroni. (2020). *Review of: "ReITopic: A Graph-Based Semantic Relatedness Measure in Topic Ontologies and Its Applicability for Topic Labeling of Old Press Articles"*. Qeios. doi:10.32388/sdj16w.