

Leveraging Open Access publishing to fight fake news

Since the very first experiences in Open Access publishing at the end of 20th century, (arXiv and PLOS, two pioneers of open access distribution of academic articles were created in 1991 and 2001, respectively), Open Access has developed tremendously.

Today, a significant fraction of research is published open access. Evaluation estimates it to be as high as 28% [Piwowar, 2018] and it occupies an ever-growing position in the scientific debate with the adoption, in 2018 of the plan S which creates an European level mandate for Open Access.

In addition to being ethically desirable per se, there are many academic, economic and societal arguments in favor of open access. These arguments, based on an improvement of the exploitation and reuse of research results, are well described theoretically in the literature [Tennant, 2017]. Nevertheless, the practical demonstration of the use of Open Access outside research communities are not common, and we have not many reports of these. The objective of our project is to illustrate the possible uses of Open Access outside of academia.

In this study, we will examine how open access combined with the right machine learning tools can help fight fake news.

Natural Language processing has been revolutionized these last years, by the use of neural networks based language models such as word2Vec [Mikolov, 2013] and Bert [Devlin, 2018].

By building space representation of the words and concepts used in texts, these models are able to take into account the meanings of studied texts. These methods have been shown to be of use to create knowledge bases from corpus of texts [Petroni, 2019] in a unsupervised manner. More specifically, [Tshitoyan, 2019] has shown that these methods, applied to a scientific corpus in an unsupervised manner, were able to retrieve the links between concepts that exists in the texts.

This study will investigate how these principles will be used to build a text-mining pipeline that indicates whether a scientific claim is backed by the scientific literature or not.

In this exploratory phase, the following methods will be applied:

- data from Euro Pubmed Central database will be used to train a Word2Vec model.
- claims will be restricted to health-related questions of the pattern “Does X cure/cause/prevent Y?”.
- Claims will then be classified by exploring the links between X, Y and the concept of cure / cause / prevent as learned in the language model.

The pipeline will be evaluated with claims taken from expert-based scientific fact-checking network such as metafact.io or sciencefeedback.co.

By validating the principle of fact-checking scientific claims with Open Access literature, we hope to pave the way to improved automatic fact-checking tools, which will allow an increased understanding of research results by the broad public and to show a strong impact of open science in society.

References

- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Retrieved from <http://arxiv.org/abs/1810.04805>
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. Retrieved from <http://arxiv.org/abs/1301.3781>
- Petroni, F., Rocktäschel, T., Lewis, P., Bakhtin, A., Wu, Y., Miller, A. H., & Riedel, S. (2019). Language Models as Knowledge Bases? Retrieved from <http://arxiv.org/abs/1909.01066>
- Piwowar, H., Priem, J., Larivière, V., Alperin, J. P., Matthias, L., Norlander, B. Haustein, S. “The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles.” *PeerJ*, 6, e4375 (2018) <https://doi.org/10.7717/peerj.4375>
- Tennant, J. P., Waldner, F., Jacques, D. C., Masuzzo, P., Collister, L. B., & Hartgerink, C. H. J. (2016). The academic, economic and societal impacts of Open Access: an evidence-based review. *F1000Research*, 5(May), 632. <https://doi.org/10.12688/f1000research.8460.3>
- Tshitoyan, V., Dagdelen, J., Weston, L., Dunn, A., Rong, Z., Kononova, O., ... Jain, A. (2019). Unsupervised word embeddings capture latent knowledge from materials science literature. *Nature*, 571(7763), 95–98. <https://doi.org/10.1038/s41586-019-1335-8>