Introduction

2021-09-13

Introduction

The amount of published research has experienced a continuous exponential growth in last decades Bornmann & Mutz (2015), and last COVID19 pandemic has further increased this trend, thanks also to a decrease in journals' processing time and to the diffusion of preprint database usage (Aviv-Reuven & Rosenfeld, 2021; Horbach, 2020; Hoy, 2020). Consequently, it gets harder for researchers and practitioners to stay up-to-date on the latest findings in their field. Secondary research is of paramount relevance in this scenario, but suffers from the same issue. The amount of scientific production to collect and screen is ever more demanding in term of time and human resources (Allen & Olkin, 1999; Borah et al., 2017, p. cohen2010evidence).
The article collection and screening phases of a systematic review are particularly problematic tasks [systematic review tutorial]. The first is the collection of published research possibly relevant to a scientific question through the an appropriately built search queries used with scientific database search engines. The construction of these search queries is a particularly delicate tasks [], that requires both domain and basic technical knowledge of each database query language to produce a set or results containing all relevant articles (high sensitivity) and as few non-relevant ones as possible (high specificity). Furthermore, often manual work is required to download the publication data from the database, store and organize them. This aspect may be complicated by limits in the number of records that can be downloaded at once [], and if multiple databases are used, by the necessity to harmonize different formats and resolve duplication []. The screening phase consists of classifying the collected publications as relevant or not for the topic of interest (Bannach-Brown et al., 2019). It is usually the more resource demanding task of a systematic review; even with appropriately built search queries, the returned results easily range in the tens of thousands and it was estimated that labelling 10 000 publications may take as much as 40 weeks of work, while the average clinical systematic review takes 63 weeks to be completed (Bannach-Brown et al., 2019; Borah et al., 2017).
The field of Data Science applied to evidence synthesis and acquisition has since years reached maturation []. Through the application of natural language processing (NLP) it is possible to transform free text into quantitative features, with various level of abstraction and generalization []; with machine learning, such text-derived features can be used to map and reproduce human judgment to automatize the screening given an initial human labeled set [].
The field of automatization of systematic reviews has been ripe of improvements in the last years [], and it is possible to foresee that such techniques will become the de facto standard [?]. Some example... Our contribution to the field is an integrated framework which tries to simplify most tasks related to systematic reviews in innovative ways: our framework provides solutions for automatic publication collection from multiple online sources and management, and a Bayesian, active machine learning based tool to support humans in the screening task, exponentially reducing the number of records to review. Finally, we propose an experimental method to semi-automatically generate efficient search queries given an already labelled set of publications.
We apply this framework in production of a systematic review evaluating the mathematical modeling of patient referral networks among hospitals and its impact on the diffusion of healthcare-associated pathogenic microrganisms with a focus on antimicrobial resistant strains []. The systematic review original protocol is published in [].

Allen, I. E., & Olkin, I. (1999). Estimating time to conduct a meta-analysis from number of citations retrieved. *Jama*, *282*(7), 634–635.

Aviv-Reuven, S., & Rosenfeld, A. (2021). Publication patterns' changes due to the COVID-19 pandemic: A longitudinal and short-term scientometric analysis. *Scientometrics*, 1–24.

Bannach-Brown, A., Przybyła, P., Thomas, J., Rice, A. S., Ananiadou, S., Liao, J., & Macleod, M. R. (2019). Machine learning algorithms for systematic review: Reducing workload in a preclinical review of animal studies and reducing human screening error. *Systematic Reviews*, *8*(1), 1–12.

Borah, R., Brown, A. W., Capers, P. L., & Kaiser, K. A. (2017). Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*, *7*(2), e012545.

Borah, R., Brown, A. W., Capers, P. L., & Kaiser, K. A. (2017). Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. *BMJ Open*, *7*(2), e012545.

Bornmann, L., & Mutz, R. (2015). Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the Association for Information Science and Technology*, *66*(11), 2215–2222.

Horbach, S. P. (2020). Pandemic publishing: Medical journals strongly speed up their publication process for COVID-19. *Quantitative Science Studies*, *1*(3), 1056–1067.

Hoy, M. B. (2020). Rise of the rxivs: How preprint servers are changing the publishing process. *Medical Reference Services Quarterly*, *39*(1), 84–89.

Larsen, P., & Von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by science citation index. *Scientometrics*, *84*(3), 575–603.