

August 23, 2024

Att. Dimitris Mavridis & Terri Pigott
Editors, *Research Synthesis Methods*

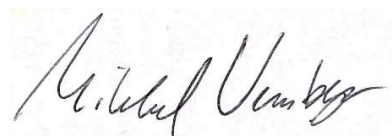
Dear Dr. Mavridis and Dr. Pigott

For your consideration, we are submitting the manuscript “GPT API Models Can Function as Highly Reliable Second Screeners of Titles and Abstracts in Systematic Reviews: A Proof of Concept and Common Guidelines” The manuscript is 42 pages long including all relevant references, tables, figures, notes, and highlights. Concretely, the article includes 3 figures. R codes to reproduce all examples and figures can be found at <https://osf.io/apdfw/> and https://github.com/MikkelVembye/screen_benchmarks. Along with the paper, we have developed the R package that supports the major aims of the article. This open-source software can be assessed at <https://mikkelvembye.github.io/AIscreenR/>. The manuscript has not previously been submitted or published elsewhere.

My co-authors and I believe that the article will be relevant for the readers of the *Research Synthesis Methods* for the following reasons. First of all, we introduce a novel method for screening titles and abstracts with OpenAI’s GPT API models and show that these models can yield screening performances on par with humans and in some cases even better than human screeners when compared to typical human screener performances in high-quality systematic reviews. To support future applications of the screening approach, we develop a novel benchmark scheme and guideline for when it is appropriate to use GPT API models of title and abstract screening in high-quality reviews. Moreover, we present the workflow for how such screenings can be standardized in high-quality reviews, and present the R package AIscreenR to support this purpose. Altogether, we hope that these contributions can ease some of the challenges encountered in independent duplicate screening, reducing rigid human labor, and ultimately making GPT API models acceptable as independent second screeners within high-quality systematic reviews

Thank you for considering this work.

Sincerely,



Mikkel Holding Vembye
Researcher
VIVE - The Danish Center for Social Science Research
Phone: (+45) 3131 9209
Mail: mihv@vive.dk