

# GPT API Models Can Function as Highly Reliable Second Screeners of Titles and Abstracts in Systematic Reviews: A Proof of Concept and Common Guidelines

Mikkel Holding Vembye<sup>1</sup>, Julian Christensen<sup>1</sup>, Anja Bondebjerg Mølgaard<sup>1</sup>, & Frederikke Lykke Witthöft Schytt<sup>1</sup>

<sup>1</sup> VIVE - The Danish Center for Social Science Research

## Author Note

Mikkel Holding Vembye, <https://orcid.org/0000-0001-9071-0724>, Department of Quantitative Methods, VIVE. Julian Christensen, <https://orcid.org/0000-0002-4596-6998>, Department of Governance and Management, VIVE. Anja Bondebjerg Mølgaard, <https://orcid.org/0000-0002-2825-4921>, Department of Quantitative Methods, VIVE. Frederikke Lykke Witthöft Schytt, Department of Quantitative Methods, VIVE. This research was funded by VIVE Campbell. All authors declare no conflict of interest.

Supplementary materials for this article are available at OSF <https://osf.io/apdfw/> and GitHub [https://github.com/MikkelVembye/screen\\_benchmarks](https://github.com/MikkelVembye/screen_benchmarks). The AIScreenR R package presented in the article can be assessed at <https://mikkelvembye.github.io/AIScreenR/>. A preprint of this article is available at <https://osf.io/preprints/osf/yrhzm>.

We would like to thank Jens Dietrichson, Trine Filges, Terri Pigott, Tiril Borge, Heather Melanie R. Ames, and Christopher James Rose for valuable comments and sharing of screening data. Also thanks to Sofie Elgaard Lisager Jensen and Johan Klejs for testing the AIScreenR software and for valuable inputs to the workflow.

Correspondence regarding this article should be addressed to Mikkel H. Vembye, Soeren Frichs Vej 36 G, 8230 Aabyhoej, Denmark. E-mail: [mihv@vive.dk](mailto:mihv@vive.dk)