



PVLDB Copyright License Form

Proceedings of the VLDB Endowment

Title of Work: Large Language Models as Data Preprocessors

Author(s): Haochen Zhang, Yuyang Dong, Chuan Xiao, and Masafumi Oyamada

Select either option A or B by checking [X] the appropriate box below:

☐ A: The undersigned certifies that the work identified above (the "Work") has been written in the course of employment by a government agency or other organization whereby, as a result of such employment, no copyright in the Work exists.

☒ B: If and when the Work is accepted by the VLDB Endowment for publication, the undersigned hereby grants to the VLDB Endowment the following:

1. An exclusive, worldwide, irrevocable and royalty-free right to publish the Work first in audio, video, electronic, and/or digital form, and any and all other media and distribution mechanisms now known or later developed;
2. A non-exclusive, worldwide, perpetual, irrevocable and royalty-free license to reproduce, use, publicly perform, modify the formatting of and create derivative works of the Work;
3. The right to license to third parties ("Third Party Publishers") the right to reproduce, use, modify the formatting of, create derivative works of the Work and to publish, publicly perform, distribute and sell copies of the Work as a compilation or collective work, for any commercial or non-commercial purpose, in digital form and all other media and distribution mechanisms now known or later developed; and
4. The right to license to Third Party Publishers the right to reproduce, use, modify the formatting of and create derivative works of excerpts of the Work and to publish, publicly perform, distribute and sell copies of such excerpts incorporated with other works of authorship as a compilation or collective work, for any commercial or non-commercial purpose, in digital form and all other media and distribution mechanisms now known or later developed; provided that such compilation or collective work shall include an acknowledgement of the author's contribution.

The license in sub-paragraph 1. above includes right to (a) publish and distribute the Work under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>) and (b) obtain a Digital Object Identifier ("DOI") for the Work.

The right in sub-paragraph 3 above includes permitting Third Party Publishers to use the DOI for the Work to enable others to access an abstract of the Work or portions of the Work and purchase digital copies of the Work.

The license and rights granted above do not transfer ownership of the copyright in the Work to the VLDB Endowment. The author(s) reserve(s) all rights in the Work not expressly granted.

This instrument must be signed by at least one author of the Work or, in the case where the Work was commissioned by another person or organization or created as part of the author's duties as an employee, an authorized representative of the commissioning person or organization or of the employer.

I am an author of the Work, an agent of an author of the Work, a representative of the person or organization that commissioned the Work or a representative of the author's employer and I represent and warrant that I have full power and authority to grant the above license of the Work.

Print Name: **Chuan Xiao**

Signature *Chuan Xiao*

Date **08/06/2024**

Title (if not author)