CDEMapper 2.0: Expediting collaborative review and facilitating consensus building for Common Data Elements mapping

Vincent J Zhang, MS, Huan He, PhD, Lingfei Qian, PhD, Yujia Zhou, MS, Yan Wang, PhD, Ruey-Ling Weng, MS, Jihoon Kim, PhD, Hua Xu, PhD, Na Hong, PhD

Department of Biomedical Informatics and Data Science & Department of Emergency Medicine, School of Medicine, Yale University

Abstract

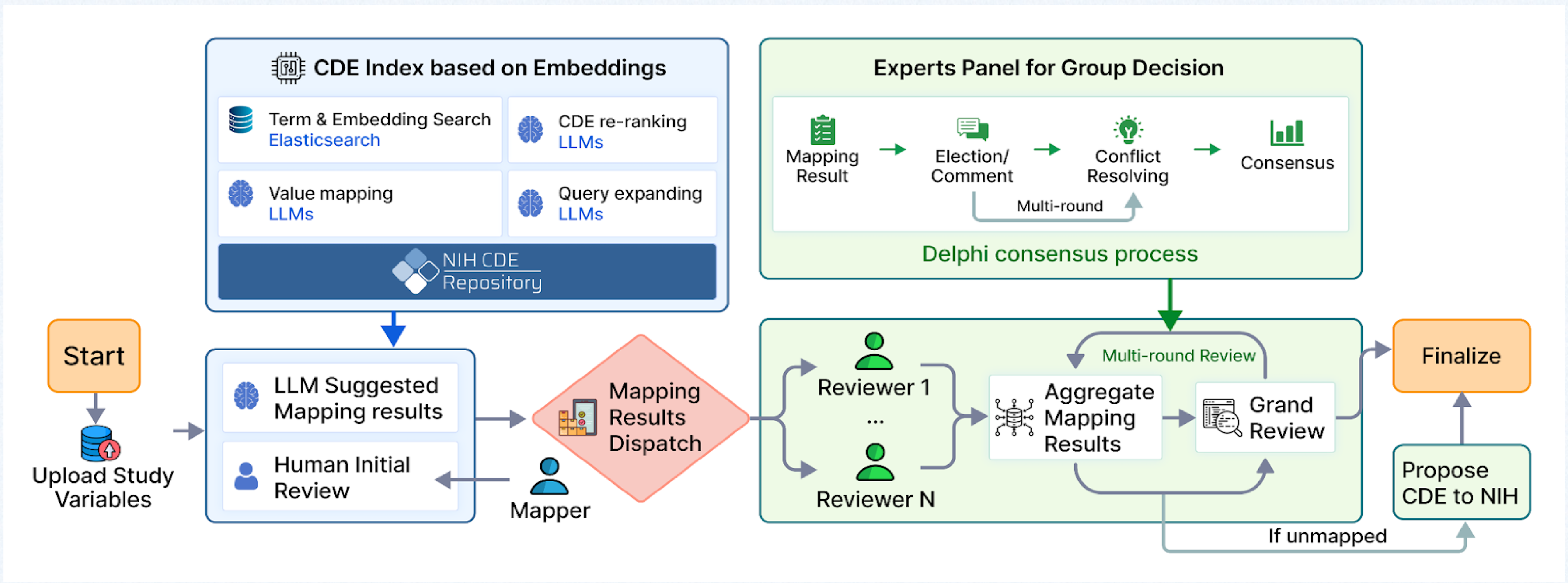
The NIH promotes the use of Common Data Elements (CDEs) to improve research data interoperability. However, aligning study variables with CDEs is complex and labor-intensive. We developed CDEMapper, an LLM-assisted tool that streamlines this process and has been expanded in version 2.0 to support collaborative expert review. It supports multi-user workflows, reduces manual effort, and facilitates consensus building. Future work will focus on broader evaluation and version control.

INFORMATICS PROFESSIONALS. LEADING THE WAY.

Introduction

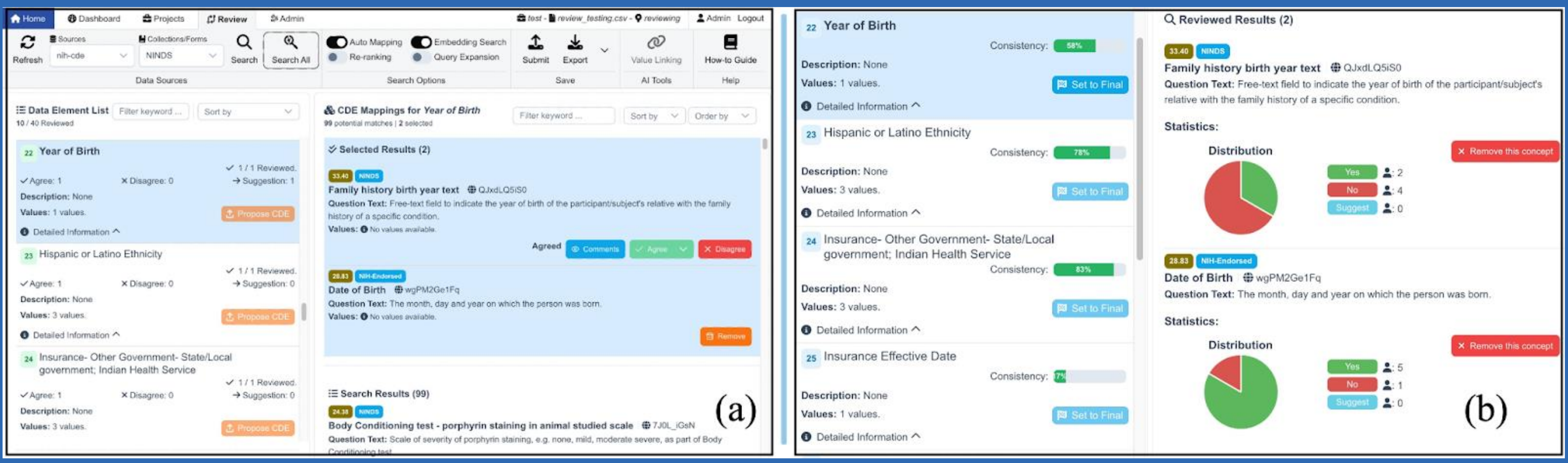
The adoption of Common Data Elements (CDEs) remains limited due to the substantial time, expertise, and effort required to align study variables with existing CDEs. To address this, we developed CDEMapper, an LLM-powered tool that streamlines CDE mapping and reduces manual workload. While automation improves efficiency, challenges persist in collaborative settings: (1) experts from different disciplines may emphasize different aspects of mapping, (2) terminology inconsistencies complicate agreement, and (3) consensus often requires iterative discussions. Existing tools provide limited support for structured collaboration. To overcome these barriers, we introduced CDEMapper 2.0, featuring multi-user review workflows, voting mechanisms, and conflict resolution tools that facilitate expert consensus and promote real-world CDE adoption.

Workflow



Guided by the Delphi methods, we support multiple role management features and enable collaborative review of mapping tasks by multiple users in CDEMapper 2.0. Domain experts can make shared decisions within the tool through iterative rounds of review, voting, refining, and commenting. The workflow involves 1) a data/project manager (Admin) initiates a mapping project and assigns corresponding roles to Mappers and Reviewers; 2) an initial mapping is conducted by a Mapper with the assistance of LLMs, generating preliminary mappings; 3) these preliminary mappings are subsequently dispatched to multiple Reviewers for independent review, voting, and feedback; 4) the review results submitted by multiple reviewers are aggregated and presented to the Admin, who compares the results, conducts grand review (organizing a panel discussion if needed), and determines whether additional review rounds are necessary; and 5) if no mappings are found in the NIH CDEs, experts can propose new CDEs within the tool.

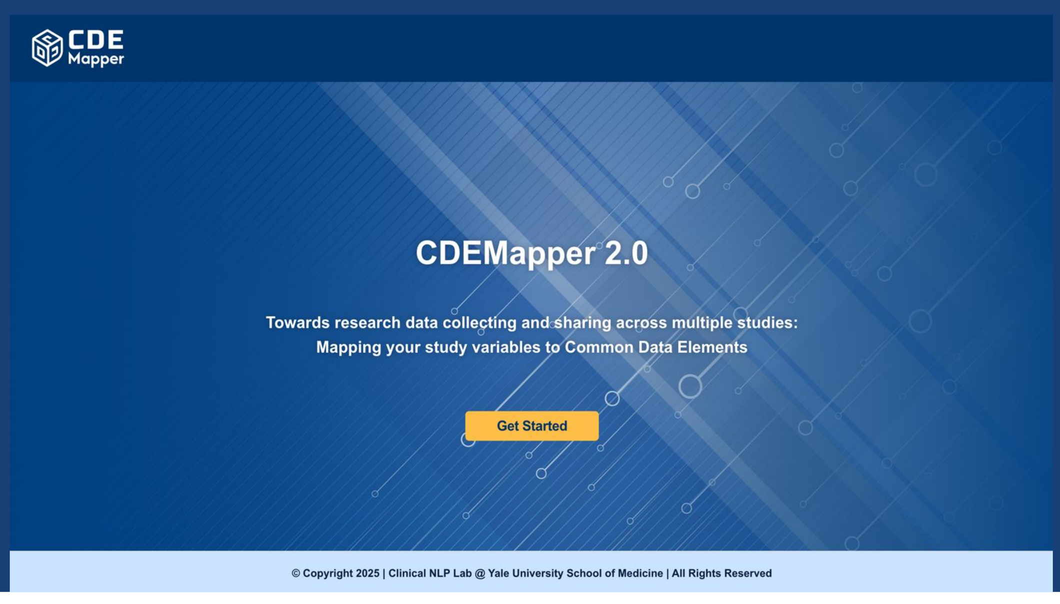
User Interface



The main user interfaces: (a) shows the reviewing interface, where reviewers can vote, comment on preliminary mappings, and explore or add new mappings; and (b) presents the grand review page, enabling consistency status checking, group-based discussion, conflict resolution, and initiation of subsequent review rounds.

Conclusion

CDEMapper 2.0 addresses a key barrier to CDE adoption by combining LLM-driven automation with structured, multi-expert review. A key contribution of this work is transforming CDE mapping and review from a largely manual, time-consuming task into an efficient, collaborative, LLM-assisted team process. Future work will focus on broader validation, expert engagement, and integrating version control to keep pace with evolving NIH CDE standards.



https://cdemapper.clinicalnlp.org

Scan QR code for entering website. For the further question please contact: Vincent.j.zhang@yale.edu