

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

Vincent J Zhang, MS<sup>1</sup>, Huan He, PhD<sup>1</sup>, Lingfei Qian, PhD<sup>1</sup>, Yujia Zhou, MS<sup>1</sup>, Yan Wang, PhD<sup>2</sup>, Ruy-Ling Weng, MS<sup>1</sup>, Jihoon Kim, PhD<sup>1</sup>, Hua Xu, PhD<sup>1</sup>, Na Hong, PhD<sup>1</sup>

<sup>1</sup>Department of Biomedical Informatics and Data Science, School of Medicine, Yale University, New Haven, CT, USA; <sup>2</sup>Department of Emergency Medicine, School of Medicine, Yale University, New Haven, CT, USA

## Introduction

The National Institutes of Health (NIH) promotes the use of Common Data Elements (CDEs) to enhance research data interoperability. However, adoption remains limited due to the significant time, expertise, and effort required to align study variables with existing CDEs. To address this, we developed CDEMapper[1], a large language model (LLM) powered tool that streamlines variables to CDE mapping and reduces manual effort. Despite improving automation, we observed challenges in collaborative settings: (1) Experts from different disciplines may prioritize different aspects of mapping, leading to potential conflicts in mapping results. (2) Terminology inconsistencies complicate agreement. (3) Mapping often requires iterative discussions and consensus. To overcome these limitations, we refactored our system into CDEMapper 2.0, which includes structured multi-user review workflows and enhanced features for conflict resolution and expert consensus formation.

## Methods

CDEMapper integrates semantic search via GPT-based embeddings and GPT-assisted re-ranking to improve CDE recommendation accuracy. In CDEMapper 2.0, we introduced a Delphi collaborative review framework with support for multi-role management, enabling domain experts to reach consensus through iterative reviewing, voting, and commenting. The workflow involves (Figure 1) (1) an Admin initiating a project and assigning roles; (2) a Mapper generating initial mappings using LLM; (3) reviewers independently reviewing and voting; (4) the Admin aggregates feedback, leading discussions if needed, and finalizing decisions, and (5) the option to propose new CDEs when necessary. This structured, expert-driven process enhances both the quality and efficiency of mapping.

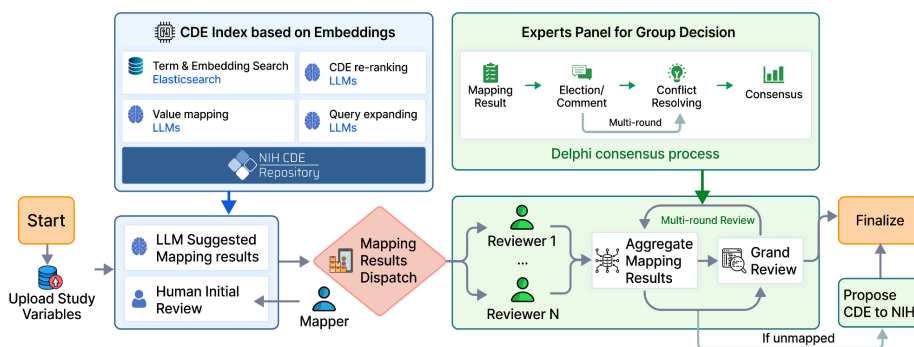


Figure 1. Overview of the LLM-assisted CDE mapping and collaborative review workflow

## Results

Table 1 shows the mean severity scores and standard deviation (SD) for each usability heuristic item. The severity scores ranged from 0.5 to 2.2 for the 10 items on the heuristic checklist. (Score interpretation: The heuristic evaluation rating score ranged from 0 to 4 based on how much an issue affects the user experience. A score of 0 indicates no usability issue, 1 indicates a cosmetic issue only, 2 indicates a minor usability issue, 3 indicates a major usability issue, and 4 indicates a usability catastrophe). The area identified as most in need of improvement was “Help and Documentation,” where users noted that the tool would benefit from more intuitive guidance and training resources.

## Discussions and Conclusions

CDEMapper 2.0 addresses a critical barrier to the adoption of CDEs by enabling LLM-assisted and collaborative review workflows. A key contribution of this work is the transformation of CDE mapping and review from a largely manual, face-to-face task into an online, collaborative, LLM-assisted, team-based process. Our future work will involve testing the tool on more datasets, engaging a broader range of domain experts to further evaluate the tool, and incorporating version control and quality assurance in terms of constantly updated NIH CDE versions.

Table 1. Usability evaluation results

Usability Heuristics	Mean (SD)
Visibility of System Status	1.0 (0.4)
Match Between the System and the Real World	1.2 (0.7)
User Control and Freedom	1.2 (0.6)
Consistency and Standards	1.3 (0.5)
Error Prevention	1.3 (0.5)
Recognition Rather than Recall	0.8 (0.2)
Flexibility and Efficiency of Use	1.3 (0.9)
Aesthetic and Minimalist Design	0.8 (0.2)
Help Users Recognize, Diagnose, and Recover from Errors	1.5 (0.4)
Help and Documentation	1.7 (0.5)

## Reference

[1] Wang Y, Huang J, He H, et al., CDEMapper: Enhancing NIH Common Data Element Normalization using Large Language Models. arXiv preprint arXiv:2412.00491. 2024. Available at: <https://arxiv.org/abs/2412.00491>.