

OCL Mapper and the CIEL Bridge: Open-Source, AI-Assisted Mapping of Clinical Concepts to Standard Vocabularies and OMOP

Jonathan Payne^a, Joseph Amlung^b, Filipe Lopes^c, Andrew S. Kanter^d

^aOpen Concept Lab; ^bRegenstrief Institute; ^cOrango I/O (Health); ^dDepartment of Biomedical Informatics, Columbia University

Background

Mapping local source concepts to standard vocabularies is the difficult, time-consuming gateway step of OMOP ETL, and it is constrained by a global shortage of trained terminologists [1]. A powerful way to ease it is to route mapping through a mature, expert-curated clinical interface terminology. The Columbia International eHealth Laboratory (CIEL) dictionary — a designated digital global good of tens of thousands of clinical concepts, curated and maintained over many years and each mapped to multiple international reference terminologies (ICD-10, ICD-11, SNOMED CT, LOINC, RxNorm) — serves as a semantic bridge to standards [2]. A local concept is mapped once to CIEL and then inherits CIEL's deterministic, expert-maintained crosswalks to many targets at once — map once, reach many — and, because CIEL is represented in Athena as a non-standard vocabulary, it also reaches standard OMOP concepts [3]. CIEL is already the de facto reference dictionary for OpenMRS and for numerous national and humanitarian health systems, so the bridge rests on terminology that is openly available and continuously curated. **OCL Mapper operationalizes the CIEL bridge at scale** and pairs it with AI assistance for the concepts CIEL does not yet cover. Built on the Open Concept Lab terminology server — a global good and the primary distribution medium for CIEL — the Mapper is open source. In a recent ICD-11 evaluation, the CIEL bridge was the single strongest and least-replaceable source of correct candidates — evidence that the bridge methodology is accurate as well as complementary to other methods.

Methods — what the demonstration shows

Attendees will see the complete mapping workflow on real data, with the CIEL bridge as its centerpiece:

1. **Project setup.** A source file of clinical terms is loaded; an ICD-11 target and a CIEL lookup source are configured.
2. **Bridge mapping + multi-algorithm candidate generation.** For each term the Mapper first seeks a CIEL match and follows CIEL's curated maps into the target terminology — the bridge — yielding deterministic, expert-vetted candidates, often including the precise postcoordinated code. In parallel it queries the WHO ICD-11 linearization search and autocode APIs and a custom "LLM-as-terminologist" ICD-11 Agent that constructs postcoordinated expressions [4]. The CIEL bridge supplies the high-precision, human-curated candidates; the other algorithms broaden coverage for the concepts CIEL has not yet mapped.
3. **AI-assisted adjudication.** The OCL AI Assistant ("LLM-as-judge"; Claude Haiku 4.5) reviews the pooled candidates and returns a primary recommendation plus up to two alternates, each classified RECOMMENDED, CONDITIONAL, INSUFFICIENT, or REJECT with a confidence signal and a short rationale.
4. **Human-in-the-loop review.** A terminologist inspects, accepts, or overrides each recommendation; candidate provenance — including whether a candidate arrived via the CIEL bridge and the CIEL concept it came from — and postcoordination are visible throughout.
5. **Export.** Reviewed mappings are exported for downstream use; because CIEL cross-walks to standard OMOP concepts, the same output feeds an OMOP ETL pipeline, complementing Usagi in a CIEL-first, Usagi-second workflow [5].

Because CIEL maps to many standards, the same bridge applies to ICD-10-WHO, LOINC, and SNOMED CT targets, among others. Figure 1 shows the overall mapping process; Figure 2 shows the candidate-review interface.

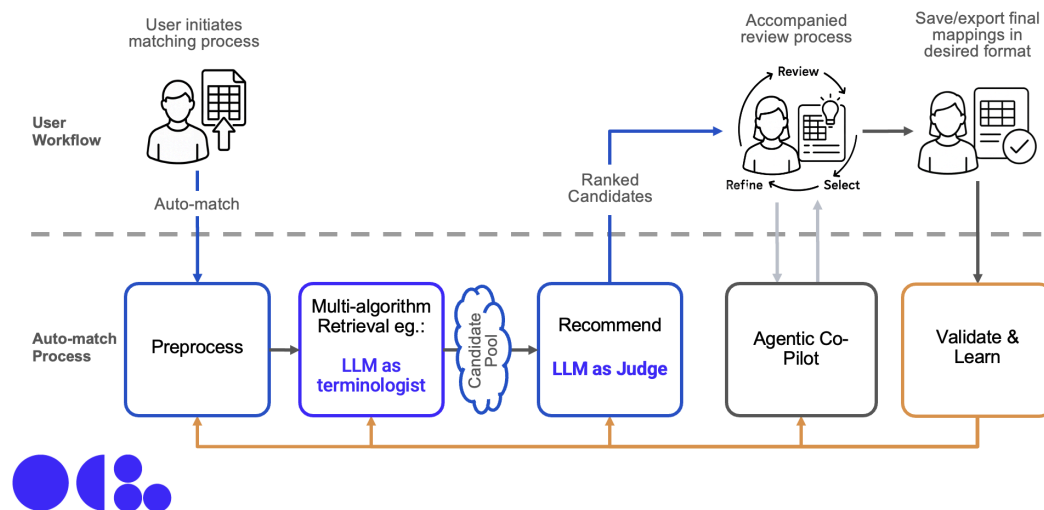


Figure 1. Overview of the OCL Mapper process, with AI components (LLM-as-terminologist and LLM-as-judge) highlighted in blue.

Results — validation and field use

The bridge methodology is supported by both evaluation and field use. In a 100-term ICD-11 validation under strict exact matching, the CIEL bridge was the strongest single source: correct at rank one for 51 of 100 terms, and the only source of the correct code for 14. With an AI judge, adding the bridge to WHO's autocode API raised exact matches from 42 to 60 (poster #105, Kanter et al.). CIEL's role is equally central in the field: it underpins national EHR dictionary mapping toward ICD-11 in Uganda and Malawi, where local concepts reach international standards through CIEL, and for OMOP ETL the OHDSI Africa working group maps source concepts to CIEL first, resolving only the remainder with Usagi [5]. More broadly, OCL Mapper has been applied by Médecins Sans Frontières across 15 programs in Mosul, Iraq, cutting clinical-form-development time by roughly 80%, and by a rural health NGO in Burundi — with no dedicated terminologist — to map more than 1,500 concepts in under two months.

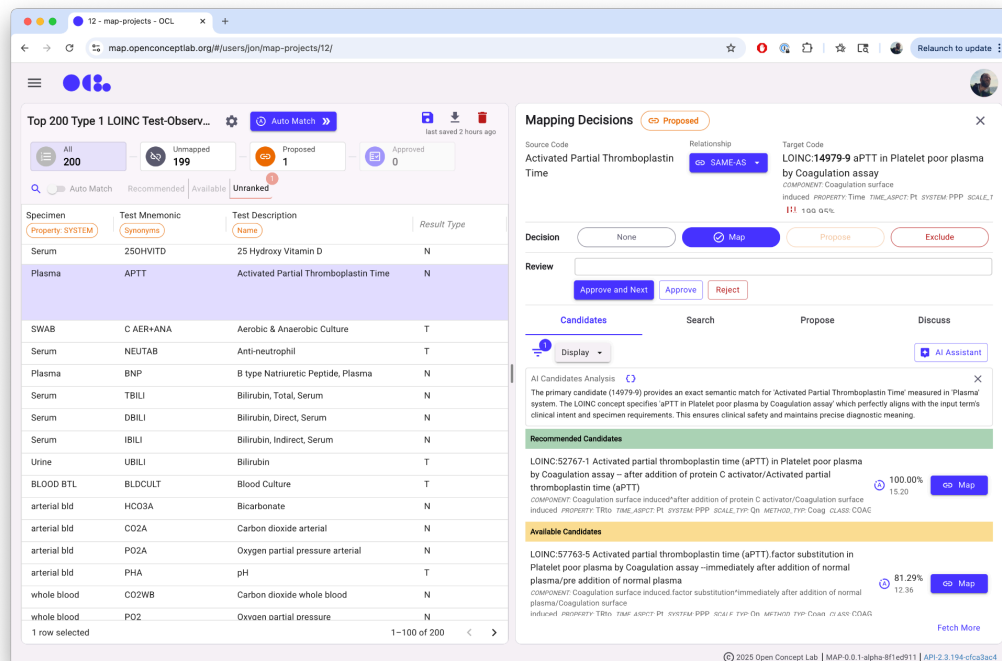


Figure 2. The OCL Mapper candidate-review interface.

Conclusions

OCL Mapper pairs the **CIEL interface-terminology bridge** with AI assistance and human review, turning a costly bottleneck into a tractable workflow. The bridge is the methodological core: CIEL contributes deterministic, expert-maintained maps that reach many standard terminologies at once — and, via Athena, standard OMOP concepts — while the AI terminologist and judge extend coverage to the concepts CIEL has not yet mapped. An accompanying evaluation (poster #105) found CIEL to be the strongest and least-replaceable single source, complemented rather than replaced by the AI components — a reminder that curated terminology remains foundational even as AI advances. Because both the tool and the CIEL dictionary are open, the OHDSI community can evaluate the approach on their own data and adopt it upstream of existing tooling, in a CIEL-first, Usagi-second workflow. We welcome collaboration, and we invite attendees to bring a handful of their own source concepts and map them live at the software demo.

Source code: github.com/OpenConceptLab/oclmapper; Mapper page: openconceptlab.org/tools/mapper; CIEL: app.v3.openconceptlab.org/orgs/CIEL

References

1. Hripcsak G, Duke JD, Shah NH, et al. Observational Health Data Sciences and Informatics (OHDSI): opportunities for observational researchers. *Stud Health Technol Inform.* 2015;216:574–578. PMID: 26262116.
2. Rosenbloom ST, Miller RA, Johnson KB, Elkin PL, Brown SH. Interface terminologies: facilitating direct entry of clinical data into electronic health record systems. *J Am Med Inform Assoc.*

2006;13(3):277–288. PMID: 16501181.

3. Observational Health Data Sciences and Informatics. The Book of OHDSI. 2021.
<https://ohdsi.github.io/TheBookOfOhdsi/>
4. Harrison JE, Weber S, Jakob R, Chute CG. ICD-11: an international classification of diseases for the twenty-first century. BMC Med Inform Decis Mak. 2021;21(Suppl 6):206.
doi:10.1186/s12911-021-01534-6.
5. Usagi: mapping source codes to OMOP standard vocabularies. OHDSI.
<https://ohdsi.github.io/Usagi/>