

From Tool Zoo to FAIRified Pipeline Evaluating Terminology and Mapping Tooling for the SHINe Platform

Pre-project work to define and evaluate requirements for collaborative terminology work – A Swiss-wide effort to strengthen the semantic interoperability

SHINe – Swiss Health Interoperability Network- cross-institutional infrastructure project for a semantic collaboration platform, Swiss Data Science Center and Inselspital Bern accepted in the 2025 Swiss Data Science Center Call

Result 1: Requirements

- Terminology work is often not harmonized between institutions
- Divergent modeling, inconsistent metadata, variable use of exchange profiles, inconsistent mapping governance
- Multiple projects share datasets and aim for aligned pipelines, yet execute terminology work with heterogeneous tooling and workflows.
- No coherent toolchain for collaborative, multi center project work of different medical domains

Result 2: Evaluation of tools

Ecosystem	Tool	Primary role for SHINe	Strengths summary	Limitations summary
SNOMED	Snap2SNOMED	Collaborative simple maps to SNOMED CT	User friendly, collaboration and roles, built-in review and versioning	Limited to SNOMED CT; import and export constraints for broader interoperability
SNOMED	Authoring Platform	Full SNOMED extension authoring	Powerful governance and workflow, MRCM-backed validation	Requires authoring access and licensing; steep learning curve; heavy for mapping-only use
SNOMED	SNOMED CT browser with MRCM view	Modeling transparency for review	Strong lookup, hierarchy, and constraint visibility	Not an authoring tool; bulk automation depends on APIs
SNOMED	Snowstorm	Terminology services API backbone	ECL and history, FHIR terminology endpoints, extension maintenance APIs	Limited out-of-the-box UI; integration work needed
SNOMED	Simplex Toolkit	Spreadsheet-based reference set / map maintenance	Practical iterative workflow, ACCEPTED-only quality gate, RF2 export	Depends on Snowstorm; can feel process-heavy for small teams
OHDSI	Athena	Vocabulary version alignment	Reliable baseline distribution, consistent OMOP vocab bundles	Manual download/update; licensing can restrict some vocabularies
OHDSI	Usagi	OMOP concept mapping	Easy UI, strong human-in-the-loop controls	String-similarity limits; manual effort remains significant
OHDSI	WhiteRabbit / Rabbit-In-a-Hat	ETL scoping and documentation	Strong shared ETL artifact, early gap detection	Not a terminology mapping tool; time cost for profiling and documentation
OHDSI	Data quality Dashboard and Achilles	Cross-site quality assurance (QA) and characterization	Transparent rule sets, scalable QA, widely adopted	Not focused on terminology authoring; triage burden can be high
OHDSI related	Carrot Mapper	Metadata-driven collaborative OMOP mapping	Strong permission model, shareable mapping rules, container-friendly deployment	OMOP-specific; depends on good source profiling; separate transform step

The need – Conclusion of pre-project work:

A modular toolchain would meet the needs – > “three layer” architecture

- (i) an API-centric terminology services (FHIR–first where possible), with central elements or the specific
- (ii) OHDSI components for OMOP vocabulary alignment, cross-mappings, mapping review, and CDM–level quality assurance, and
- (iii) SNOMED–centric components for extension lifecycle management and MRCM–constrained validation.

Feedback loop between mapping data and source annotation quality. Relying on hospital–integrated tools. User–friendly interface to share mappings easily. Strong user management and quality assurance. Curated mappings can be exported back to improve hospital data at the source. Project specific versioning. Concept repository. Create a sustainable system for continuous improvement.

Aim of project: A collaborative “Terminology & Mapping Platform” that supports governance, validation, versioning, and reproducible releases

- Promote practical, consensus–oriented approaches to terminology management and semantic alignment
- Provide open, reusable OMOP–aligned mappings and semantic artifacts that support consistent and interoperable data representation across Swiss institutions
- Complement the foundations established by national as well as local semantic initiatives

Karen Triep¹, Stefan Milosavljevic², Hugo Armando Guillen–Ramirez³, Laure Vancauwenberghe², Marcel Messerli⁴, Almut Lütge², Cyril Matthey–Doret², Christophe Gaudet–Blavignac⁵, Julien Ehrsam⁶, Jean Louis Raisaro⁷, Solange Zoergiebel⁸, Oksana Riba Grognez², Olga Endrich^{1,9}



¹ Inselspital, University Hospital of Bern; Health Information Management, Medical Directorate; Bern, Switzerland, ² Swiss Data Science Center (SDSC), Swiss Institute of Technology, Zürich/Lausanne, Switzerland, ³ Department for BioMedical Research (DBMR), University of Bern; Bern, Switzerland, ⁴ Inselspital, external consultant, Bern University Hospital, Bern, Switzerland, ⁵ Division of Medical Information Sciences University Hospitals of Geneva, University of Geneva, ⁶ SIMED, Division of Medical Information Sciences University Hospitals of Geneva, Geneva, Switzerland, ⁷ Biomedical Data Science Center, Centre hospitalier universitaire vaudois, Lausanne, Switzerland, ⁸ IT Data Science & Management, Centre hospitalier universitaire vaudois, Lausanne, Switzerland, ⁹ Insel Data Science Center; Inselspital, University Hospital of Bern; Bern, Switzerland