Optimizing Multilingual Terminology Resources for Specialized Translation: Methodological, Pedagogical and Technological Perspectives from the Cochrane Project*

Marija Omazić[1,*,†], Mihovil Gazda[1,†], Karla Malović[1,†] and Sven Kroflin[1,†]

[1] *Josip Juraj Strossmayer University of Osijek, Lorenza Jaegera 9, Osijek 31000, Croatia*

Abstract

This paper presents a case study on the optimization of an English-Croatian digital terminology resource used in the translation of Cochrane Plain Language Summaries (PLSs). Focusing on a real-world volunteer and trainee translation workflow set up by Cochrane Croatia, the study examines how terminological data quality, resource structure, and user-oriented design affect usability, consistency, and translator decision-making in a high-risk medical domain. The analysis is based on a redesigned Cochrane termbase extracted from the Phrase CAT tool. It identifies recurrent terminological and structural problems, including heterogeneous entry types, unmarked synonymy, and insufficient usage instructions. The optimization process was guided in part by international terminology standards and by trainee translator needs, resulting in a resource that better supports trainee translators who are not domain experts, while maintaining terminological coherence. This paper presents the conceptual and technical development of a terminology-support application that integrates terminology management, standardized phraseology, and an interactive style guide within a pedagogical framework to be used in a terminology course. Our *MediTerm* application has been further developed through a relational database model, a backend service, a REST API, and a frontend prototype, moving the project from a conceptual proposal toward a viable digital resource for translator training and specialized translation use. In this study we argue that terminology resources in such contexts can function not only as professional reference tools but also as learning tools, and that systematic, user-oriented terminology optimization can enhance both translation quality and translator training.

Keywords

terminology management, multilingual terminology resources, specialized translation, medical translation Cochrane, terminological data quality, translation pedagogy

1. Introduction

In translation practice today, multilingual digital terminology resources have become indispensable rather than optional reference tools. Their role is particularly important in specialized domains such as medicine, law, and technology, where terminological precision is directly linked to communicative success, ethical responsibility, and institutional credibility [1]. Terminology resources affect translator decision-making, constrain term variation, and contribute to textual coherence across large translation projects that are carried out by multiple translators.

Medical translation is a high-risk domain in which terminological inaccuracies may have far-reaching consequences. When medical information is communicated to lay audiences, as is the case with Cochrane Plain Language Summaries (PLSs; see https://www.cochrane.org/hr/evidence), translators must sometimes choose between scientific accuracy and user accessibility. Terminology resources in such cases function as mediators between expert knowledge and public understanding of complex medical conditions.

* *5th International Conference on "Multilingual digital terminology today. Design, representation formats and management systems" (MDTT) 2026, June 25-26, 2026, Zadar, Croatia.*

[1,*] Corresponding author.

[†] These authors contributed equally.

momazic@ffos.hr (M. Omazić); mgazda@ffos.hr (M. Gazda); kmalovic@ffos.hr (K. Malović); skroflin@ffos.hr (S. Kroflin)

0000-0001-6383-166X (M. Omazić)

This paper builds on and substantially expands the findings of the MA thesis by Dikšić [2], who investigated the optimization of English-Croatian terminology and translation instructions for volunteers that are used in the Cochrane translation workflow.

2. Cochrane *Plain Language Summaries*

Cochrane is an international non-profit organization that produces, maintains, and disseminates systematic reviews of healthcare interventions. A central part of its medical knowledge dissemination strategy are Plain Language Summaries, which are a condensed text genre designed to communicate key findings to lay audiences without specialist medical training [3]. Plain language summaries have a number of distinctive linguistic and terminological features. They rely on controlled terminology, avoidance of unnecessary technical jargon, explicit definition of key concepts, and they should be generally reader-oriented. Translating such texts requires linguistic competence, solid understanding of the medical field and terminological mediation [4], as translators must ensure consistency while adapting terminology to the expectations and background knowledge of the target audience.

The ongoing cooperation between the Faculty of Humanities and Social Sciences (FFOS), University of Osijek, and Cochrane Croatia started in 2022, and it plays an important role in the training of volunteer and trainee translators using real-life assignments in a real-life setting. This cooperation is embedded in the real-world application of medical translation and it connects academic training and professional practice. Volunteer translators involved in this are predominantly MA students of translation studies who take the elective course *Introduction to Terminology*. The translation workflow is run by the Croatian Cochrane coordinator at the Faculty of Medicine, University of Split. The translations are done in the Phrase CAT tool, using a Cochrane-provided termbase co-created by users and verified by the main Cochrane coordinator, as well as written translation instructions. All volunteer translations done as graded course assignments are checked multiple times before publication: by a peer volunteer student translator, by a language editor (course instructor), and finally by a medical expert (the Cochrane coordinator). The resources made available to translators serve as normative guides, effectively defining what is considered an acceptable or preferred translation solution, but in the process of translation, we have encountered many issues due to their structure and quality.

3. From user needs to resource design: terminological and structural optimization

User-oriented terminology theory emphasizes that terminology resources must be tailored to the needs, competences, and working conditions of their intended users [1, 5]. Standard terminology design, as reflected in ISO 704 [6] and the broader tradition from Wüster onward, implicitly assumes a competent domain specialist who requires fast, unambiguous retrieval, which has recently been efficiently challenged as practically unattainable [7]. Translating for Cochrane complicates this assumption considerably. Volunteer and trainee translators differ significantly from expert professional translators in terms of domain knowledge, terminological experience, and tolerance for ambiguity. The features that make a terminology resource efficient for experienced professional users are not always compatible with the needs of beginners. A professional translator consulting a termbase needs a single preferred equivalent, no redundant explanation and immediate retrieval. A trainee translator using the same resource needs something closer to what in translation theory would be called scaffolding [8], i.e. explicit reasoning behind preference decisions, visible alternatives with indications of why they are preferred or deprecated, and enough contextual guidance to support independent decision-making over time. Terminology resources can serve both professional and pedagogical function, as indicated in literature on translator training, as terminological competence is a learnable skill that resources can actively scaffold rather

than merely support [9, 10], which is consistent with arguments for scaffolding and situated learning in translation education [8].

The analysis conducted by Dikšić [2] showed that when these requirements are not met, users must compensate through doing additional external research, which increases cognitive load and undermines terminological consistency across the project. The original Cochrane termbase was not designed with these user needs sufficiently in mind. Its heterogeneous structure, lack of preference marking, and inclusion of non-terminological material made it difficult to use efficiently, particularly for less experienced trainee translators.

The first version of the termbase consisted of 559 entries stored in a single Excel spreadsheet exported from Phrase. From the terminological perspective, it conflated multiple conceptual levels, including basic terms, multi-word terms, fixed phrases, combined with occasional informal instructions to translators, which violated fundamental principles of terminological data modelling [1, 6]. The optimization carried out by Dikšić [2] involved a comprehensive cleaning and redesign. Terminological units were separated from phraseological material, grammatical forms were normalized, duplicates were eliminated, alphabetical ordering was enforced, and synonym clusters were consolidated, with preferred terms explicitly marked.

These changes were guided in part by terminological theory and in part by personal observation and reports of actual user difficulties. The resulting consolidated resource was more usable for both professional and trainee users, but the problem described above was not fully resolved. A termbase, however well designed, is a static repository. It can indicate what the preferred term is, but it cannot readily explain why, model the decision-making process, or adapt its level of explicitness to the user's level of expertise. This motivated us to develop the terminology-support application MediTerm proposed in Section 5 of this paper. Three modules of that application, a terminology module, a phraseology module, and an interactive style guide, operate as linked but discrete components. This means that an experienced translator can use the terminology module alone for rapid retrieval, while a trainee translator is offered more “scaffolding” and the fuller resource and can progressively learn the required norms.

4. Terminological data quality, standards and usability effects

Terminological data quality is a central determinant of the reliability and effectiveness of terminology resources [11]. In the analysis, Dikšić identified five major categories of issues in the original termbase: technical errors, orthographic inconsistencies, lack of clarity, excess information, and grammatical problems.

Technical errors included inconsistent capitalization of basic terms such as *patient* and *protocol*, which may signal unintended semantic distinctions. Orthographic inconsistencies involved non-standard spelling and inconsistent use of Croatian diacritics. Lack of clarity was most evident in synonym clusters, where multiple Croatian equivalents were provided without any indication of preference or context. Excess information manifested in the inclusion of explanatory phrases within term entries, while grammatical problems involved incorrect inflectional forms. Systematic correction of these issues significantly improved terminological coherence and user trust in the resource. As all users of Cochrane Phrase could add terms to the termbase, this has led to entries that are not medical terms at all (like conjunctions *and*, *but*, or names of months like *May*), and occasionally completely wrong entries, which undermined the trustworthiness of the resource. These entries were eliminated in the optimization process, but remained in the CAT tool termbase nonetheless, despite us trying to argue for uploading the corrected version.

The optimization she carried out was guided by international terminology standards, particularly ISO 704 and ISO 1087:2019 [6, 12], which emphasize univocity, consistency, transparency, and user orientation. While the revised resource does not claim full standard compliance, its design decisions align with standard-recommended best practices. At the same time, the study highlights the limitations of spreadsheet-based terminology management. Such formats offer limited support for metadata enrichment, versioning, validation workflows, and

collaborative editing. These limitations motivate the exploration of more dynamic, pedagogically informed terminology management solutions, which is what we attempted to do in the further optimization steps presented in this paper.

These improvements, however, also revealed a more fundamental limitation, not specific to the Cochrane termbase but inherent to terminology stored in CAT tools generally. CAT tool termbases are architecturally optimized for fast in-segment lookup during active translation. This design focus limits the functional scope of terminology resources and how they can be used [13]. Specifically, CAT-embedded termbases typically operate on a flat data model that offers limited or no support for layered metadata such as register tags, validation status, deprecation marking, or hierarchical conceptual relations. Cross-referencing between terminological data, phraseological patterns, and stylistic guidance is architecturally outside the scope of what a CAT termbase is designed to do. This, in combination with the reluctance of the project managers to update the CAT termbase, has initially yielded a corrected spreadsheet version to be consulted by the subsequent cohort of students, which further complicated their workflow. As they could not trust the client CAT termbase, they had to consult the spreadsheet version, which slowed them down and created even more tension and uncertainty.

For a translation context involving trainee translators working in a high-stakes medical domain, these constraints are not trivial. A trainee who encounters a deprecated synonym, an unmarked register distinction, or a term without usage context has no resource within the CAT environment to resolve the uncertainty, and is therefore likely to either make an uninformed choice or abandon the resource and conduct external research, which undermines the purpose of having a curated termbase. The spreadsheet format to which the original termbase had been exported from Phrase compounded these limitations further. Spreadsheets offer no support for versioning, collaborative validation, or the kind of hierarchical and relational structuring that terminology theory, and in particular ISO 704, recommends as standard for concept-oriented termbase design.

These structural constraints, which persisted even after data quality optimization, provided the primary motivation for the terminology-support application we propose here in Section 5. The application is explicitly designed as a complement to the existing CAT environment, preserving the value of in-segment lookup during active translation, while providing a richer, pedagogically informed resource for the terminology work that precedes and informs the translation process itself.

5. Conceptual design of a terminology-support application *MediTerm*

Building on the empirical findings of the thesis by Dikšić [2], this paper proposes the development of a dedicated terminology-support application by 2024/2025 cohort students Mihovil Gazda, Karla Malović and Sven Kroflin, tentatively called *MediTerm – Master Medical Terminology* (see Figure 1 below). The application is conceived as a hybrid pedagogical and professional tool, designed to complement the existing CAT environment rather than replacing it. The main objective of the application is to make terminological knowledge explicit, structured, and learnable. Unlike traditional termbases, which often function as static repositories, the proposed application emphasizes explanation, contextualization, and transparency of terminological decision-making. The application consists of three integrated, interoperable modules (components of the application) that communicate in real-time. It includes a terminology module, a phraseology module, and an interactive style guide.

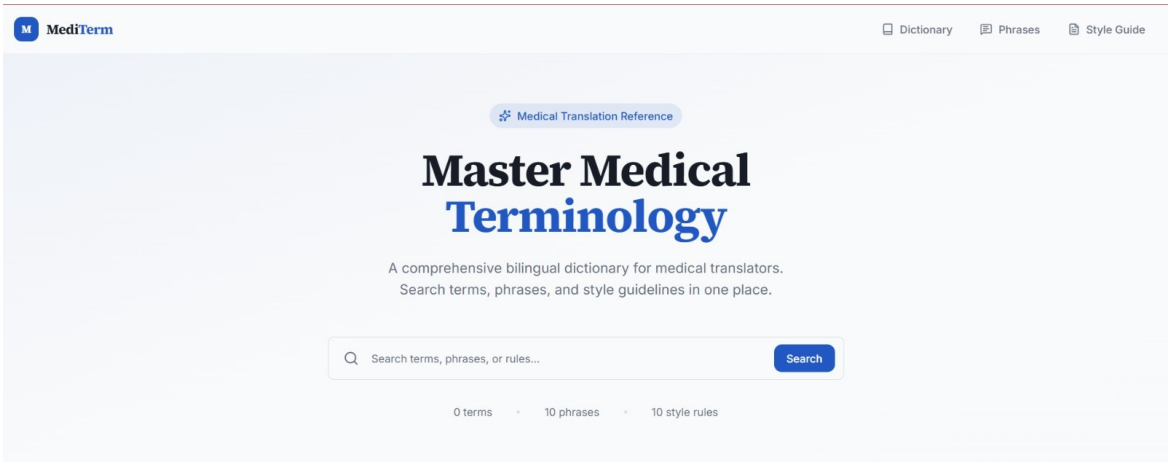

Figure 1: Screenshot of the MediTerm application user interface

The terminology module (contextual dictionary module, see Figure 2 below) is the core module of the application. It goes beyond simple lexical retrieval by addressing a critical gap in current medical translation tools, which is the distinction between professional and lay registers. This feature is particularly relevant for the translation of Cochrane Plain Language Summaries (PLSs), in which patient health literacy is of greatest importance. The system maps medical concepts to dual registers. For instance, a query for the term *myocardial infarction* would yield two distinct, register-tagged results: *infarkt miokarda* (tagged as Professional/SPC), intended for medical professionals, and *srčani udar* (tagged as Layman/PIL), intended for patient-facing materials. The application clearly indicates which translation is which, and the two terms are linked. Furthermore, the dictionary module includes real-time validation status (*preferred, admitted, deprecated*), discouraging the use of obsolete or incorrect terminology through visual cues. Usage recommendations are indicated for each term so the application discourages the use of terms that are no longer current. This is enabled by a complex network of synonyms and related terms in the database. Beyond validation and register, the module includes a hierarchy structure. The application recognizes hyperonymy and hyponymy, providing users with a comprehensive view of the terminological hierarchy, where applicable. By mapping terms to their superordinate and subordinate concepts, the application enables effortless navigation between different levels of terminology. Additionally, through Contextual Intelligence, the application integrates real usage examples of terms drawn from medical databases. Contextual Intelligence here refers to the system's ability to retrieve and display authenticated usage examples from curated medical corpora (Cochrane, HALMED, PhD thesis repositories) alongside each entry.

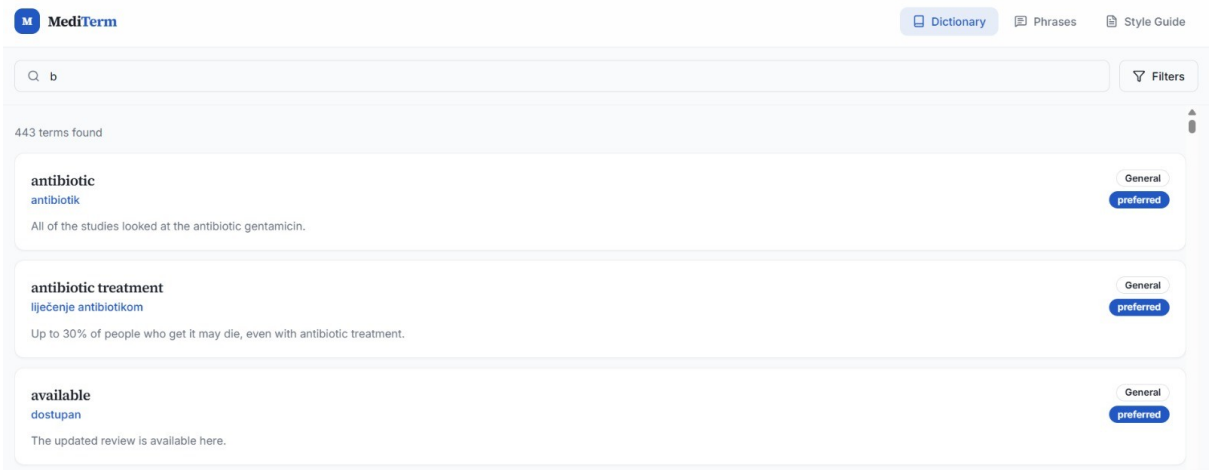

Figure 2: Screenshot of the MediTerm Dictionary module

The phraseology module (see Figure 3 below) is based on an extensive corpus of standardized Cochrane texts. Serving as a repository of standardized syntactic blocks, the purpose of this module is to ensure terminological and stylistic consistency on a broader level, not only within a single text, but across the entire Cochrane database in Croatian. This module functions as a boilerplate sentence repository and addresses the fact that Cochrane texts, as well as many other medical texts, contain numerous headings, titles, and sentences that appear in almost every piece of writing (e.g. *What did we do?*, *The aim of the review*, *Key findings,* etc.). To ensure stylistic consistency throughout the database, this module provides approved translations for these phrases and collocations. Unlike the first module, which focuses on individual terms, this module organizes content according to the IMRaD structure (Introduction, Methods, Results, Discussion) typical of scientific abstracts. For instance, a translator struggling to formulate the results section can filter for "*Main Results*" and retrieve approved boilerplate phrases such as "*The main findings of this review…*" (*Ključni rezultati ovog sustavnog pregleda…*) or "*Their findings are summarized below*" (*Rezultati ovog Cochraneovog sustavnog pregleda prikazani su u nastavku*). The key feature of the phraseology module is its integration with the terminology module. If the translator searches for *adverse effects* in the terminology module, they will not only receive the full range of information described above but will also be encouraged to follow the link to the phraseology module to find common sentences in which this term appears, thereby saving time and ensuring stylistic consistency.

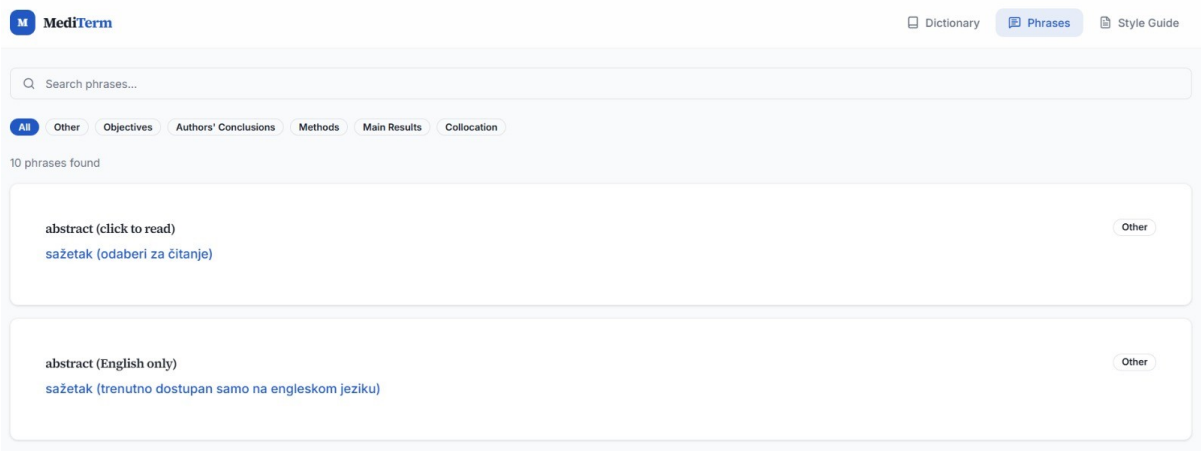

Figure 3: Screenshot of the MediTerm Phrases module

The final module (see Figure 4 below), interactive style guide module, functions as a knowledge base that further enhances the application by providing users with an interactive, browsable, and searchable style guide for text translation. Instead of lengthy theoretical explanations, users are presented with a rule title, a brief description, and a concrete side-by-side example of correct and incorrect usage, which are visually distinguishable. This module is contextually linked to the terminology module. Users can access the style guide through links from various terms in the terminology module that relate to specific rules. For instance, if a user searches for a term such as *West Nile fever* in the dictionary module, and is unsure how to capitalize such a term, the system provides a direct link to the style guide module, more specifically, to the relevant rule (e.g. the Orthography and Grammar category, Capitalization segment). The user can then identify the specific example or a similar one and apply the rule, ensuring the final text is grammatically and stylistically polished. Additionally, the module serves as a structured knowledge base, allowing users to access the module independently of the translation process and use it as a handbook. The content is divided into several categories. The Translation Aligned with Croatian Linguistic Conventions (Prevođenje u duhu hrvatskoga jezika) category offers style guidelines such as avoiding anglicisms, word order rules, and avoiding literal translations. The Ortography and Grammar (Pravopis i gramatika) category focuses on the norm and includes rules on capitalization,

comma placement, declension, use of passive and active voice, prepositions, etc. The Semantic Nuances (Semantika) category addresses issues such as "false friend" translation, translation of modal verbs, and nuances of word meaning. The Technical Instructions (Tehničke upute u prevođenju) category provides text formatting rules, instructions on copying tags in CAT tools, and tips on using the first and second module to ensure accuracy and style consistency. Finally, this module includes dedicated sections for Most Common Mistakes (Najčešće pogreške u prevođenju) and Best Practices (Savjeti za prevoditelje i lektore).

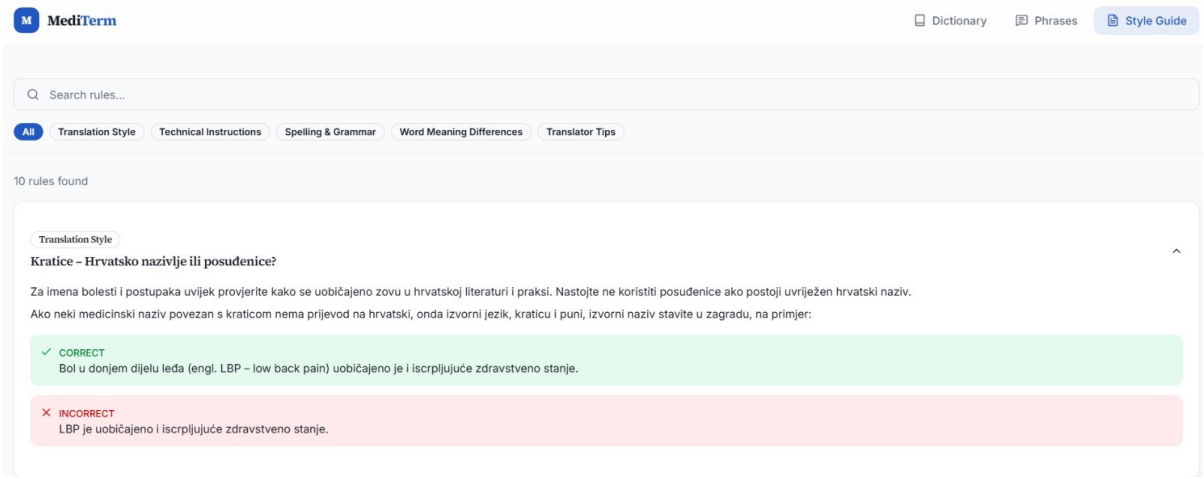

Figure 4: Screenshot of the MediTerm Style Guide module

To illustrate the pedagogical potential of the application, a standardized workflow scenario is proposed where a translator or a student is working on a Cochrane Plain Language Summary. In the initial search phase, the user encounters the source phrase "*The study reported serious adverse events*" and is unsure whether to translate adverse event as *štetni događaj* or *nuspojava*. They search for *adverse* in the terminology module, and the system responds with several results: *adverse event, adverse effect*, and *adverse reaction*. The user selects the *adverse event* result, sees the correct translation, and avoids a critical error. Additionally, the system offers the user a list of style guide rules where they can learn more about the difference. Following this, during register selection, the user encounters the term *myocardial infarction* and looks it up in the terminology module. They receive two options tagged with their specific registers: *infarkt miokarda* (tagged as Professional/SPC) and *srčani udar* (tagged as Layman/PIL). The user recognizes the difference, and since the target text is a plain language summary, they use *srčani udar*. In the following phase of contextualization and standardized phrasing, the user has the correct terms but struggles to translate a sentence or a subtitle. They switch to the phraseology module, and instead of searching for a word, they use a category filter (e.g., Main Results). The application shows a list of approved phrases, and the user finds the English phrase "*The main findings of this review suggest that…*" alongside the Croatian "*Ključni rezultati ovog sustavnog pregleda upućuju na to da…*". The user then combines the term from the previous phase with this phrase to produce a correct, standardized sentence. Finally, the user is unsure about the appropriate use of passive versus active voice constructions. They switch to the style guide module and select the Orthography and Grammar category. In this category, they find the title regarding passive usage, where the application provides a brief explanation of the rule. The user identifies the relevant syntactic pattern and applies the rule.

Although still in development for the English-Croatian language pair, the underlying database architecture is adaptable, allowing for future scalability to other languages without significant codebase modifications.

6. Prototype implementation

In the most recent development phase, the application has been implemented as an initial technical prototype. The architecture was designed to reflect the three-part structure of the resource: terminology, phraseology, and style guidance. The database was organised around the following core entities: source_concepts, translations, phraseology, and style_rules, with additional relational tables representing links between concepts, translations, and style rules. This structure enables the modelling of one-to-many relations, such as one source concept having multiple target-language translations, as well as recursive relations representing hierarchical links between broader and narrower concepts.

The prototype was developed through the analysis of an existing Cochrane termbase, various terminology sources, and translation instructions. In addition, new terminological and phraseological data was gathered from Cochrane materials, medical terminology sources, and examples encountered during the workflow. The collected material was then examined in terms of entry type, terminological status, register, synonymy, conceptual hierarchy, and usability. Special attention was given to grouping synonymous and near-synonymous terms, as well as to modelling hyperonymic and hyponymic relations between broader and narrower concepts. This made it possible to distinguish between terms that represent the same concept, terms that are conceptually related, and terms that belong to different levels of the same terminological hierarchy.

The current database model implements the terminological principles discussed earlier in the paper. The source_concepts table stores English source terms, domains, usage examples, register information, synonym-group data, and hierarchical parent-child relationships. The translations table contains Croatian equivalents, language codes, register labels, validation status, didactic notes, usage examples, and links to relevant style rules. The phraseology component is represented by a dedicated phraseology table, which includes English phrases, approved Croatian equivalents, language codes, phrase categories, and links to relevant concepts. The style_rules table stores the rule category, title, explanatory content, source-language examples, correct Croatian solutions, and incorrect examples.

This relational structure is significant because it prevents the application from functioning as a flat bilingual word list. Instead, it enables the system to retrieve and display terminological, phraseological, and didactic information as interconnected layers of knowledge. For example, a single term can be linked to its Croatian translation, register information, validation status, related concepts, phraseological patterns, and relevant style rules. The database design therefore directly supports the pedagogical aim of the application: users are not only given an equivalent of a concept, but are also guided towards understanding why a particular equivalent is preferred in a particular context.

The backend was implemented using the Spring Boot framework, following a layered architecture comprising repository, service, and controller components. The repository layer manages communication with the PostgreSQL database, the service layer maps raw database data into data transfer objects for the client side, and global exception handling ensures clearer responses in the event of errors. A RESTful API was developed to enable the frontend to retrieve terminology entries, phraseological units, and style-guide rules via standard HTTP requests. Communication between the backend and frontend uses JSON, while CORS configuration restricts access to authorised client-side sources.

The prototype also introduced several practical development considerations relevant to terminology-resource design. First, the database was prepared for future multilingual expansion by including language-code fields and category identifiers, allowing additional language pairs to be added primarily through data population rather than major structural changes. Second, the use of controlled categories for domains, registers, phrase types, and style-rule categories supports more consistent filtering and retrieval. Third, the process of importing and testing data revealed typical data-quality issues, such as character-encoding problems with Croatian diacritics and incorrect synonym linking. These issues confirm that terminology-resource development requires linguistic

and conceptual modelling, technical validation, import testing, and iterative correction of relational links.

The development process revealed several challenges of building a comprehensive terminology database. The first challenge was the transformation of a spreadsheet-based terminology resource into a relational database capable of supporting synonym relations, hierarchical links, register labels, phraseological connections, and style-rule references. This required a shift from a flat list of equivalents to an interconnected knowledge model.

A second challenge concerned data consistency. Problems such as unmarked synonymy, inconsistent category values, and insufficient usage guidance became apparent during implementation because they directly affected what the user would see in the application. Testing also revealed incorrectly linked synonym relations, confirming that terminology-resource quality control must include both linguistic revision and technical validation of database relations.

A third challenge involved the modelling of register. Since the application supports both expert-oriented and patient-oriented medical communication, a single source concept may need to be linked to several Croatian equivalents, each with its own register, validation status, notes, and examples. This is especially important in the translation of Plain Language Summaries, where translators must balance accuracy with accessibility for lay readers.

Overall, these challenges showed that terminology-resource development is not only a process of collecting and storing equivalents, but an intricate task of linguistic, conceptual, and technical modelling.

Usability is increasingly recognized as a critical factor in the effectiveness of terminology resources [5], and it is best evaluated qualitatively through task-based observation and reflective feedback from translators. The initial usability evaluation was conducted across two cohorts of translation studies MA students participating in the Cochrane volunteer translation workflow: the 2024/2025 and 2025/2026 cohorts, with 20 participants in total. All participants were enrolled in the *Introduction to Terminology* course and had active experience translating Cochrane PLSs using the original resources plus the optimized spreadsheet database. Participants consistently reported that the original resources did not meet basic terminological standards. The most frequently cited issues aligned closely with the five data quality categories we identified in the formal analysis: entries were perceived as structurally heterogeneous, with no clear distinction between terminological units, fixed phrases, and translator instructions; synonym clusters were unresolved and unhelpfully large; and preference marking was largely missing, leaving participants uncertain which equivalent to use in a given context. In addition to these structural observations, participants identified several usability problems with direct consequences for translation quality. Occasional contradictory information in the termbase and inconsistencies between the termbase and other available Cochrane resources created confusion during active translation tasks. Conflicting Croatian translations of recurrent PLS headings, precisely the type of content addressed by the proposed phraseology module, were reported as a source of confusion, uncertainty and inconsistency. Difficulties in choosing between synonyms, particularly in the absence of register guidance, were described as requiring additional external research that disrupted the workflow and increased task completion time. These findings confirm that the usability problems observed were not just cosmetic but had measurable consequences for translator decision-making and consistency. Participants' intuitions about what would improve the resource, separating terms from phrases, normalizing grammatical forms, enforcing ordering, and providing clearer preference and register guidance, correspond directly to the design principles that informed the termbase optimization and the architecture of the proposed application. This alignment between user-reported needs and our design decisions strengthens the ecological validity of the proposed application.

A more comprehensive usability study will be conducted with the 2026/2027 student cohort once the fully functional application is available for testing. It will consist of a task-based protocol with discrete, pre-defined tasks, including term lookup, synonym selection, register identification, and phrase retrieval. Quantitative indicators will include average lookup time, number of external searches performed per task, and error rate. Qualitative data will be collected through reflective

feedback and follow-up interviews. This design will allow direct comparison with baseline performance data from the current cohort and provide an empirical basis for evaluating the application's effects on translation quality, terminological consistency, and cognitive load in a medical translation training.

7. Conclusion

Our Cochrane case study shows that terminology optimization is a multidimensional intervention that can encompass data quality, usability, and translator training pedagogy. It illustrates how terminology resources can also be prepared to function as learning tools. The application we are developing adopts this perspective by including terminological knowledge into a learner-centered digital environment. This reflects broader trends in digital humanities, open science, and competence-based translator education. Our study shows that systematic terminology optimization has the potential to improve multilingual translation workflows, terminological data quality, but also pedagogical outcomes and quality of translation. By combining terminological theory, empirical analysis, and pedagogical design, the Cochrane case provides a transferable model for other domains involving volunteer or trainee translators.

But the implications of this work go beyond methodology and pedagogy. Medical translation for lay audiences has an ethically significant position. A terminological error in a Cochrane Plain Language Summary reaches patients and caregivers who may act on the information they receive. In a workflow involving trainee and volunteer translators who, by definition, operate at the boundary of their competence, the quality of the terminology resource is not merely a usability concern, but a risk mitigation measure. The validation status indicators, deprecation marking, and register tagging built into the proposed application are, in this light, not only pedagogical features but also mechanisms for reducing the probability of consequential terminological error reaching a published text and the highly sensitive audience. Responsibility for translation quality in institutional volunteer workflows is distributed across multiple agents, the translator, the peer reviewer, the language editor, and the medical expert, and we are aware that no single resource can substitute for human verification. A well-designed terminology resource, however, can reduce the frequency with which that chain is called upon to catch errors that adequate resource design could have prevented at source.

Future work should focus on the full implementation and population of the database, usability testing with the 2026/2027 cohort, and empirical evaluation of the application's effects on translation quality and terminological consistency.

Declaration on Generative AI

During the preparation of this work, generative AI tools (ChatGPT 5.2) were used for language editing, grammar and spelling check. All content was critically reviewed, verified, and finalized by the authors, who take full responsibility for the publication’s content.

References

[1] M. T. Cabré, Terminology: Theory, Methods and Applications, John Benjamins Publishing Company, Amsterdam/Philadelphia, 1999. doi:10.1075/tlrp.1.

[2] A. Dikšić, Optimizing English to Croatian Terminology for Cochrane’s Plain Language Summaries Translation, Unpublished Master’s Thesis, Josip Juraj Strossmayer University of Osijek, Osijek, 2024.

[3] Cochrane. Plain Language Summaries: Making health evidence accessible. Cochrane Handbook, 2020.

[4] W. Karwacka, Medical translation, in: Ł. Bogucki, S. Goźdź-Roszkowski, P. Stalmaszczyk (Eds.), Ways to Translation, Wydawnictwo Uniwersytetu Łódzkiego, Łódź, 2015, pp. 271-298.

[5] H. Suonuuti, Guide to Terminology. Tekniikan Sanastokeskus, Helsinki, 2001.

[6] International Organization for Standardization. ISO 704:2022. Terminology work — Principles and methods, 2022. URL: https://www.iso.org/standard/79077.html.
[7] R. Temmerman, Towards New Ways of Terminology Description: The sociocognitive approach, John Benjamins Publishing Company, Amsterdam/Philadelphia, 2000. doi:10.1075/tlrp.3.
[8] E. Calvo, Scaffolding translation skills through situated training approaches: Progressive and reflective methods, The Interpreter and Translator Trainer 9 (2015) 306-322. doi:10.1080/1750399X.2015.1103107.
[9] M. Fernández-Parra, Terminology Management for Translators. Routledge, London, 2023.
[10] S. Montero Martínez, P. Faber, Terminological competence in translation, Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication 15 (2009) 88-104. doi:10.1075/term.15.1.05mon.
[11] K.-D. Schmitz, Terminology and localization: in H. J. Kockaert; F. Steurs (Eds.), Handbook of Terminology: Volume 1. John Benjamins Publishing Company, Amsterdam/Philadelphia, 2015, pp. 452-464. doi:10.1075/hot.1.ter7.
[12] International Organization for Standardization. ISO 1087:2019. Terminology work — Vocabulary, 2019. URL: https://www.iso.org/standard/62330.html.
[13] L. Bowker, Terminology and translation: in H. J. Kockaert; F. Steurs (Eds.), Handbook of Terminology: Volume 1. John Benjamins Publishing Company, Amsterdam/Philadelphia, 2015, pp. 304-323. doi:10.1075/hot.1.ter5.