Procesamiento del Lenguaje Natural, Revista nº 74, marzo de 2025, pp. 335-348 recibido 05-12-2024 revisado 22-01-2025 aceptado 03-02-2025

Lexico-Semantic Approaches to Terminology Mapping: A Survey

Revisión de enfoques lexico-semánticos para el mapeo terminológico

Lucía Palacios, Patricia Martín-Chozas,
Elena Montiel-Ponsoda, Pablo Calleja
Universidad Politécnica de Madrid
{lucia.palacios, patricia.martin, elena.montiel, p.calleja}@upm.es

Abstract: This survey addresses the challenges of ambiguity and term variation in terminology mapping, a critical task for achieving semantic interoperability across resources within the Linguistic Linked Data paradigm. An analysis of state-of-the-art techniques is presented, including manual, computational, and hybrid approaches. Scenarios for terminology mapping are identified: (i) identical terms with similar definitions, (ii) shared terms with differing definitions, and (iii) distinct terms with highly similar definitions, and the implications of ambiguity and term variation across domains are discussed. Finally, we propose future research directions to improve current methodologies, including approaches based on Large Language Models.
Keywords: terminology, lexico-semantic ambiguity, terminology mapping, term variation.

Resumen: Este trabajo aborda los retos que plantean la ambigüedad y la variación terminológica en la tarea de mapeo de terminologías, tarea fundamental en la interoperabilidad semántica entre recursos terminológicos en el marco del paradigma de los Datos Lingüísticos Enlazados. Presentamos un análisis detallado de las técnicas más avanzadas, incluidos enfoques manuales, computacionales e híbridos, en distintos escenarios: (i) términos idénticos con definiciones similares, (ii) términos iguales con definiciones distintas, y (iii) términos diferentes con definiciones muy similares, y además, se analizan las implicaciones de la ambigüedad y la variación terminológica en distintos dominios de especialidad. Por último, se proponen futuras líneas de investigación, como los métodos basados en Grandes Modelos de Lenguaje, para mejorar las metodologías actuales de mapeo terminológico.
Palabras clave: terminología, ambigüedad léxico-semántica, mapeo terminológico, variación terminológica.

1 Introduction

Linguistic resources play a crucial role in the development of efficient techniques for Natural Language Processing (NLP) applications (Besançon et al., 2010; Baylor, Ploeger, and Bjerva, 2023). High-quality structured resources have consistently proven essential for enabling systems to accurately process and comprehend human language (Navigli and Velardi, 2010).

These resources are characterized by their heterogeneity: each linguistic resource adopts its own format, structure, and representation (Vrandečić and Krötzsch, 2014). To address this diversity, the Linked Data (LD) paradigm was introduced, establishing practices and principles for publishing, sharing, and interconnecting information on the Web (Bizer, Heath, and Berners-Lee, 2009). In the linguistic domain, this approach is known as Linguistic Linked Data (LLD).

Heterogeneity is particularly significant in the case of terminologies, which often adopt different formats and subject schemas. These inconsistencies hinder semantic interoperability, highlighting the need for LLD to integrate resources and unify shared concepts (McCulloch, Shiri, and Nicholson, 2005).

The integration and linking of terminological resources is essential to unlocking the full potential of specialized linguistic data. Resources such as thesaurus, terminologies, taxonomies, and glossaries compile termino-

ISSN 1135-5948 DOI 10.26342/2025-74-23

logical information using various formats and standards. Connecting these resources can significantly enhance their applicability and coverage, enabling them to complement and enrich each other in meaningful ways.

This contribution reviews the main challenges, techniques, and future directions in *terminology mapping*. In the next section, we characterize terminologies and the terminology mapping problem, and in Section 3, we present the adopted methodology for conducting the review in a systematic way. Section 4 delves into those strategies for generating correspondences between terminological resources. Section 5 presents new research directions based on Large Language Models (LLMs) for the terminology mapping task. Section 6 presents the discussion that was triggered in this review along with the conclusions.

2 Characterizing Terminology Mapping

This section presents the theoretical framework that contextualizes the problem in this review. Section 2.1 sets the basis for the terminology concept and presents three scenarios of terminological ambiguity and term variation. Then, Section 2.2 defines the problem of terminology mapping in the NLP field.

2.1 Terminology definition

Terminology is a polysemous term with three main meanings: first, the name given to the discipline that deals with specialized terms; second, the set of practices that analyze and collect terms in terminological resources (also known as Terminography); and third, the resulting resources themselves (Cabré, 1993). Terminology is closely related to Lexicology (and Lexicography). While the former focuses on studying terms in an area of knowledge and creating terminological resources, the latter is concerned with the lexicon of a language and the compilation of general language dictionaries.

As is the case with dictionaries, terminological resources can also vary in nature. They may be encyclopedic, learner-oriented, or etymological, depending on the intended purpose or target audience, terminological resources can similarly take different forms. For instance, some terminological glossaries may consist solely of a list of recommended terms with their definitions. In contrast, others may be more comprehensive, span various knowledge domains, provide equivalents in multiple languages, and include synonyms, definitions, and usage notes.

Unlike dictionaries, where the starting point is a list of lexical entries with defined meanings (from a semasiological perspective), the creation of terminologies adopts an onomasiological approach (Cabré, 1993). This means that terminographers first identify a set of domain-specific concepts and their corresponding designations, providing specialized definitions restricted to that particular domain. In this regard, it can be argued that terms are typically monoreferential — that is, the relationship between a form and its meaning is unique.

However, this is not always the case. More often than not, terms exhibit phenomena typically associated with lexical items. For instance, terms can have multiple meanings, not only across different domains, but also within the same area of knowledge (referred to as homonyms). Additionally, a concept can be represented by multiple terms within the same domain or across closely related domains (known as synonyms or term variants). As a result, challenges often arise when attempting to establish semantic correspondences between terms in different terminologies.

Another aspect to consider is that terms can appear in various types of term-based resources, such as glossaries, taxonomies, thesauri, etc. Even when dealing with terms from the same area of knowledge, these resources may be intended for different purposes and vary in content and granularity, which complicates the process of finding correspondences.

After careful consideration, we devise the following scenarios for establishing a correspondence between term A in terminology 1 and term B in terminology 2, both of which belong to the same or highly related domain: a) term A and term B are identical and their concept definitions coincide or are highly similar as well; b) term A and term B are the same, but their definitions differ; c) term A and term B differ, but their definitions are highly coincident or identical.

In scenario a), we are probably dealing with the same term and the same conceptualization of the domain. In scenario b), the term is the same, but the concept is different. We may be dealing with a case of terminological polysemy due to the existence of different conceptualizations of the same domain, a phenomenon known as multidimensionality (Bowker, 1997; León-Araúz, 2017). Depending on the changes undergone by the specialized domain in question, the semantics of the concept might have changed as well. Finally, scenario c), represents a case of synonymy, also known as term or denominative variation (Freixa Aymerich, 2003; Freixa and Fernandez-Silva, 2017), which may also be reflected in the semantics of the concept to a lesser or greater extent.

2.2 Terminology Mapping task

Within the literature, several tasks have been devoted to the generation of correspondences between heterogeneous resources, such as *ontology matching*, which refers to the process of automatically determining the semantic equivalences between the concepts of two ontologies (Euzenat and Shvaiko, 2013) and *entity linking*, that is the process of aligning ambiguous mentions of entities from text to a knowledge base (Kolitsas, Ganea, and Hofmann, 2018; Hertling and Paulheim, 2023). The common requirement across these strategies is the need for semantic disambiguation to ensure accurate alignment.

The task of terminology mapping specifically focuses on generating equivalent, conceptual, and hierarchical relationships between terms from different schemas (Doerr, 2001). In this context, we refer to mappings as those relationships between terms in different resources (World Health Organization, 2021). This task also needs semantic disambiguation to ensure accurate alignment but differs from entity linking as it is exclusively based on terms (common entities), and not named entities. In the case of ontology matching, it usually considers valuable domain information present in the ontological structure (Hertling and Paulheim, 2023) to establish links. In contrast, terminology mapping relies on the lexical and semantic information contained in term entries or derived from neighboring term labels (usually, broader, narrower, and related terms). This lack of domain information in the structure makes terminology mapping distinct from ontology-based approaches. In addition, ontology matching and entity linking tasks can occur in contexts that involve general-purpose data, as opposed to terminology mapping, which specifically focuses on specialized domains.

This paper proposes a review of the literature on techniques that tackle ambiguity and term variation in terminology mapping approaches. We seek not only to document existing methodologies and tools but also to establish a theoretical basis for developing new solutions to optimize this process.

3 Methodology

The review method used in this article (described in detail in the following subsections) is inspired by works for developing systematic reviews (Dybå, Dingsøyr, and Hanssen, 2007) and other relevant studies in similar fields (Navigli, 2009; Bevilacqua et al., 2021). In line with the methodologies applied in these works, we formulate the following specific research questions that are answered throughout this review.

- **Q1**: What challenges arise from ambiguity and term variation in terminology mapping, and how are they being addressed?
- **Q2**: What are the key techniques and resources for addressing ambiguity and term variation in terminology mapping tasks?
- **Q3**: What are the future directions for resolving ambiguity and term variation in terminology mapping?

3.1 Selection of bibliographic sources

To align with the recommendations of the Dybå, Dingsøyr, and Hanssen (2007) methodology for conducting a systematic review, we aimed to answer the research questions by selecting appropriate information sources and implementing an unbiased search strategy. The research focused on different types of publications, including conference and workshop proceedings, journal papers, books, chapters, and technical documents. The following databases and bibliographic sources were used for the article search: IEEE Xplore, ACM Digital Library, SpringerLink, Semantic Scholar, and Google Scholar. The selection of these databases was justified based on their relevance and comprehensive coverage of research in the areas of linked data, disambiguation techniques, and terminologies. Only articles published in indexed journals or presented at well-established conferences such as ISWC, ESWC, TACL, COLING, and ELRA were considered for inclusion. Technical documents presenting good practices or guides were only included if they were validated by community standards such as ISO or WHO-FIC.

The second stage also includes the selection of relevant literature from those bibliographic sources, which was performed by generating queries that combine different terms from the research area (such as *terminology mapping*, *disambiguation*, *terminologies* or *linguistic linked data*). Together with these terms, the queries contained the specific linguistic information involved in the disambiguation process (e.g. *sense inventory* or *sentence embedding*), a domain of knowledge (e.g. *legal* or *medical*) and external resources or approaches used (e.g. *WordNet*, *BabelNet* or *rule-based*).

3.2 Selection of papers

Based on the initial set of collected references, a title-based filtering process was conducted, selecting papers based on the following criteria: 1) if the title is related to the specific task, 2) if the title includes an approach or methodology, 3) if the title mentions terminological resources or any reference to their integration and enrichment, and 4) if the title mentions any new project or initiative on terminology mapping.

To ensure the accuracy of this initial selection, a second manual filtering process was conducted by analyzing the abstracts, during which new keywords were incrementally identified and added as they were encountered during the review. Exceptions were made for foundational papers on disambiguation like Navigli (2009) or classic works in other fields, such as linguistics like in the case of Cabré (1993), even if their titles or abstracts did not strictly meet the selection criteria.

3.3 Methodology of analysis

Since the specific literature on terminology mapping is limited, our review also includes works on *word sense disambiguation* (WSD), that is the process of selecting the correct meaning of a word based on its context (Bevilacqua et al., 2021). Additionally, we include insights from ontology matching, particularly when linguistic resources are used, given that this task often relies on approaches and techniques relevant to term disambiguation.

To the best of our knowledge, no other survey addresses the disambiguation task within the context of terminology mapping. For this reason, we have combined various classifications to propose a framework of different approaches. Our classification is inspired by previous works from which we have extracted different categories: 1) Nikiema, Jouhet, and Mougin (2017), which includes (i) manual approaches, (ii) morphosyntactic approaches, (iii) approaches based on semantic information, and (iv) approaches using a third terminology resource as background knowledge; 2) World Health Organization (2021) which includes (i)manual approaches; (ii) computational approaches and (ii) hybrid approaches; and 3) Bevilacqua et al. (2021), which includes the category of external resources.

4 Terminology Mapping Approaches

This section is divided according to the identified categories in the previous section, grouping the reviewed works by type of approach. In addition, Table 1 is provided summarizing all analyzed works focused on the terminology mapping task, including their sources and the objectives of the linked terms, to offer a comprehensive overview of the approaches.

4.1 Manual approaches

The challenge of mapping terminologies has been studied for several decades, with early projects like *The Multilingual Access to Subjects* (MACS) (Clavel-Merrin, 2004) using manual strategies to address integration issues. Initially, the mapping was entirely manual, relying on domain experts to establish semantic relationships between terms in different terminologies. ISO standards, such as ISO/TR 12300:2014 (Health informatics. ISO/TC 215, 2014), point out the importance of human expertise in resolving ambiguities and validating equivalence. Examples of manually linked terminologies include NHS Digital's SNOMED CT to ICD-10 mappings (National Health Service, n.d.) and ERIC to LCSH mappings (Vizine-Goetz et al., 2006).

More recently, some systems have been introduced to assist professionals in performing manual mapping (Mate et al., 2021). These tools aim to streamline workflows while maintaining the rigorous evaluation necessary for high-quality mappings.

Author	Source	Target	Resources	Domain
Manual approaches				
Clavel-Merrin, 2004	LCSH, RAMEAU, SWD	LCSH, RAMEAU, SWD	-	Library Science
NHS	NHS Digital's SNOMED CT	ICD-10	-	Healthcare
Vizine-Goetz et al., 2006	ERIC	LCSH	-	Education
Formal-based approaches				
SEMIC Support Center, 2014	NALS system	MARC list US Library of Congress	Silk (WordNet)	Government Data
Bulla et al., 2022	Italian Vocabulary of Artworks	Getty Thesaurus	ArCO	Cultural Heritage
Wang et al., 2012	ICPC-2	SNOMED CT	UMLS	Healthcare
Fung and Xu, 2015	CORE Problem List Subset (7 datasets)	CORE Problem List Subset (1 dataset)	-	Healthcare
Stroganov et al., 2022	UK Biobank Clinical Codes	ICD and SNOMED CT	-	Healthcare
Rule-based approaches				
Lingren et al., 2016	EHRs	DSM-IV + UMLS (manual mapping)	-	Psychiatry
WHO-FIC	ICD-10-AM	ICD-10	-	Healthcare
Semantic-based approaches				
Qamar and Rector, 2006	openEHR Archetypes	SNOMED CT	UMLS	Healthcare
McInnes and Peterson, 2013	MSH-WSD 203 extracted terms	UMLS	-	Healthcare
Aouicha et al., 2016	WordNet, Wikipedia Category Graph (WCG) and MeSH	WordNet, Wikipedia Category Graph and MeSH	Wiktionary, Wikipedia	Healthcare and General
Hybrid approaches				
Ballatore et al., 2014	Geospatial vocabulary	GeoNames	WordNet	Geography
Fung et al., 2007	SNOMED CT	ICD-9-CM	UMLS	Healthcare
Di Franco et al., 2013	Pollution and Health Terminologies	EARTh	-	Environmental Science
Chen et al.,2018	Geographic Terminologies	Geographic Terminologies	Wordnet, Hownet, Thesaurus for Geographic Senses	Geography
Schmidt et al., 2018	IATE Criminal Terminology	SUMO	Wordnet	Legal
Dechadon et al., 2020	ILO Terminologies	ILO Terminologies	Eurovoc	Labour Law
Margarida Ramos, 2020	CorkCorpus	OntoCork	-	Materials Science
Kim et al., 2020	SNUH	OMOP CDM	SNOMED CT	Healthcare

Tabla 1: Summary of the reviewed approaches in this survey.

However, manual mappings, which rely on human experts for disambiguation through contextual analysis and domain knowledge, have several limitations. Subjective interpretation can introduce inconsistencies or biases in term mapping. It is also time-consuming and resource-intensive, which limits scalability, especially when working with large datasets.

4.2 Computational approaches

The following section highlights the state-of-the-art approaches addressing ambiguity and term variation through different techniques and using diverse data types. Each approach offers specific solutions for capturing and disambiguating the meaning of terms, depending on their domain and context.

4.2.1 Formal-based information methods

One of the most widely used approaches relies on formal-based methods. These methods focus on identifying structural and lexical similarities between terms to establish correspondences. Their main technique is based on *string-matching*, which involves finding strings that partially or completely match a pattern (Alkhamaiseh and ALShagarin, 2014) and other similar computational strategies to identify equivalent or closely related entries across terminological resources.

The approach presented in Wang et al. (2012) is based on the assumption that most terminologies share lexical similarities in their vocabularies, as they rely on the same natural languages to describe similar concepts. Additionally, the study identifies four primary lexical mapping techniques: normalized string matching, expanded term matching, substring matching, and WordNet-based matching.

Silk (Volz et al., 2009) is an open-source framework for integrating heterogeneous data sources and has been used to generate links[1] in the NALS system[2] of the Publications Office of the EU and the MARC list of the US Library of Congress[3]. Silk relies on string-matching techniques, offering linking possibilities between terms with the same lexical form, which must then be manually validated. A more recent example is the integration of the Italian Vocabulary of Artworks with the Getty Thesaurus[4]. This integration combines string-matching techniques with other NLP approaches and external resources, as the former alone is insufficient for resolving ambiguous terms (Bulla et al., 2022)

In the case of medical terminologies, where terms are associated with numerical codes, string-matching techniques alone can produce effective results. The reason behind this is that there are no ambiguous terms when each is tagged with a unique and exact concept identifier. This approach, known as the direct mapping method, does not necessarily require manual review or additional NLP modules (Zeng and Chan, 2004). The mapping between ICPC-2[5] and SNOMED CT (SCT)[6] is developed by leveraging the Concept Unique Identifiers (CUIs) shared between these terminologies via the UMLS (Wang et al., 2012). The UMLS system is organized around concepts, and one of its primary purposes is to connect different names for the same concept across various vocabularies. Similar terms in other vocabularies are implicitly linked through a shared unique concept identifier. In the work of Wang et al. (2012), the goal is to identify terms that share a common CUI in the UMLS using exact string-matching techniques. Other studies employing this method include Fung and Xu (2015) and Stroganov et al. (2022), among others.

However, these methodologies are not fully automated, as they require human intervention to validate the generated links or the integration of additional resources. Furthermore, these techniques are not always effective due to their difficulties in addressing cases involving synonyms and term variants, which complicate the accurate mapping of terms belonging to the same concept (Amin et al., 2010). Additionally, in non-medical domains, systems based on systematic terminological resources like UMLS, which rely on CUIs, cannot be applied. In these fields, there are no standardized resources to assign unique identifiers to concepts and their associated terms, requiring meanings to be inferred from context. This makes semantic disambiguation a significantly more complex task.

4.2.2 Rule-based methods

Rule-based approaches involve the development and use of computerized algorithms designed to replicate the reasoning a human would apply when generating mappings between two terminologies (as outlined in the CAT technical document (World Health Organization, 2021)).

Unlike the string-matching methods mentioned earlier, rule-based methods can also incorporate various types of information about the terms to determine whether they belong to the same concept and, therefore, share meaning. The mapping process from

[1] `https://interoperable-europe.ec.europa.eu/collection/semic-support-centre/document/tutorial-use-silk-aligning-controlled-vocabularies`

[2] `https://op.europa.eu/en/web/eu-vocabularies/authority-tables`

[3] `https://www.loc.gov/marc/bibliographic/`

[4] `https://www.getty.edu/research/tools/vocabularies`

[5] `https://www.who.int/standards/classifications/other-classifications/international-classification-of-primary-care`

[6] `https://www.snomed.org/`

ICD-10-AM to ICD-10[7] takes into account not only CUIs but also the descriptors of the codes, the hierarchical structure of the classification systems, and the similarity score (World Health Organization, 2021). These elements were used to define the rules for the computerized algorithm. Human-generated maps served as the gold standard for verifying the computerized mapping approach.

Both the creation of rules and the verification and validation of candidate links are human-driven tasks. Feedback obtained during the verification process can be leveraged to further refine the rule-based algorithm. While this approach can be efficient and computationally inexpensive, it has significant limitations in handling term ambiguity. It is most effective when working with rigid and unambiguous identifiers, such as CUIs. However, rules may fail when terms show variations, such as synonyms, typographical errors, or term variants (e.g., *hypertension* vs. *high blood pressure*) (World Health Organization, 2021).

4.2.3 Semantic information-based methods

In the process of accurately identifying the meaning of terms to establish correspondences among them, it is essential to consider every element of semantic information associated: both the semantic context in which the term is situated and the intrinsic information of the terminology itself. By *semantic context* we refer to the set of data that describes the environment of the term, including its linguistic context, associated labels, domain labels, available translations, explanatory notes, definitions and neighboring terms.

This method is contextual, as it takes into account the environment of the term, unlike the methods described above, which rely exclusively on the lexical or formal information of the terms to be disambiguated (Qamar et al., 2006). For instance, McInnes and Pedersen (2013) explored the comparison of UMLS concepts with ambiguous terms from biomedical texts using both similarity and relatedness metrics. By considering the context in which a term appears, their approach attempts to link it to one of the UMLS concepts. For example, the term *cold* could refer to the temperature (C0009264) or to the common illness (C0009443), depending on the surrounding context.

To work with the semantic information of terms, embeddings are employed, generated either through static or dynamic models. These embeddings are projected onto a vectorial space, where they can be compared using similarity or relatedness measures. Similarity measures quantify the degree of association between two words (e.g. *scissors* and *paper*). Relatedness measures encompass similarity but extend to quantify the degree to which two concepts are related within a broader context, such as their position in an is-a hierarchy (e.g. *car* and *vehicle*). Terms that surpass a defined threshold are considered matches with sufficient semantic similarity or relatedness, indicating that they belong to the same concept.

Similarly, the work presented in Kang et al. (2021) automatically maps OMOP-CDM[8] codes to SNOMED codes using embedding-based semantic similarity metrics. By implementing *deep learning* techniques, they outperform traditional lexical matching algorithms by incorporating contextual and semantic information into the mapping process.

The limitations of semantic information-based approaches come from their dependence on context; unusual or insufficient contexts can result in errors. Similarly, embedding models present challenges: static models treat words outside their dictionary as out-of-vocabulary (OOV), while dynamic models may rely on pre-training corpora that are inadequate for handling term variations or ambiguities in specialized terminologies. In addition, determining the appropriate threshold for considering two terms as similar or related is often complex and subjective.

4.3 External resources-based methods

When terms lack sufficient semantic context within terminologies, it becomes necessary to rely on external sources. These knowledge sources are diverse and can exist in both structured and unstructured formats. In this section, we provide a brief analysis based on the type of resource and the nature of its format, following the classification outlined in previous related surveys (Navigli, 2009). Additionally, these external resources can be

[7] `https://www.ihacpa.gov.au/health-care/classification/icd-10-amachiacs`

[8] `https://ohdsi.github.io/CommonDataModel/`

combined with any of the previously discussed methods to support disambiguation, as will be explored in Section 4.4.

1. Computational lexicons-based: These resources can be defined as language resources from the general domain that are normally hierarchically structured. They include not only semantic information but also grammatical, morphological, phonetic, and syntactic details. Some relevant examples are WordNet[9] and BabelNet[10]. The fully unsupervised disambiguation approach using WordNet, as proposed by Ben Aouicha, Hadj Taieb, and Ben Hamadou (2016) disambiguates polysemous terms based on surrounding words, using feature vectors derived from WordNet synsets, including taxonomic information. Another notable proposal is the one developed by Ballatore, Bertolotto, and Wilson (2014). The authors propose to link heterogeneous terminologies via intermediate connections to WordNet, extracting relevant synsets based on criteria like frequency, overlap with term definitions, and manual taxonomy selection. The need for domain expert knowledge indicates a key limitation: while resources like WordNet are useful for general linguistic tasks, they often lack the granularity required for terminological resources(Bevilacqua et al., 2021).

2. Ontology-based: Ontologies serve as frameworks that semantically organize the properties, relationships, and categories of general or domain-specific data. A well-known example is UMLS (Unified Medical Language System), while being primarily a thesaurus, also functions as an ontology by structuring biomedical concepts. UMLS facilitates the linking of more specific terminologies, aiding in the integration and disambiguation of terms in specialized contexts. For example, in the medical domain, Fung et al. (2007) used UMLS to map terms from SNOMED CT to ICD9CM, using it as a knowledge base to generate semantic and lexical links. Moving to a different field, Schmidt et al. (2020) linked the IATE[11] criminal terminology to the general-purpose SUMO ontology, achieving more precise and reusable terminology alignments. Many works encounter issues with abstraction and granularity, as highlighted by Schmidt et al. (2020). For example, in their study, terms such as *kidnapping* from the criminal terminology were linked to the SUMO top-level concept *UnilateralGetting*.

3. Machine-readable dictionaries-based: Traditional resources like the Collins English Dictionary[12] and the Oxford Dictionary of Contemporary English[13] have been digitized and are accessible online. These are lexicographic resources designed to be processed by computer systems, allowing efficient access to rich linguistic information (Navigli, 2009). Wiktionary[14], a free, multilingual, open-access dictionary, stands out for its collaborative nature, offering definitions, etymology, pronunciation, translations, and usage examples. It is valuable for WSD due to its contextual examples and semantic links and is used for terminology alignment tasks thanks to tools like Wiktionary Matcher (Portisch, Hladik, and Paulheim, 2019). However, its effectiveness for terminology mapping is limited by inconsistencies in term granularity, subject coverage, and the lack of specialized domain vocabulary. Additionally, its reliance on community contributions can result in discrepancies in accuracy.

4. Thesauri-based: Thesauri can be defined as structured resources that organize terms hierarchically and thematically, facilitating search and analysis in specialized fields (Kilgarriff and Yallop, 2000). Specifically, the structured design of thesaurus makes them effective tools for addressing ambiguity by providing detailed semantic context. Prominent examples like EuroVoc[15], UNESCO[16], and AgroVoc[17], which are adapted to the Semantic Web, enable alignment with other linguistic resources, enhancing their interoperability. Dechandon, Gerencsér, and Ruiz (2020) leveraged EuroVoc as a central thesaurus for the semantic integration of terminologies from the ILO (International Labour Organization) Thesaurus. Furthermore, Chen,

[9] `https://wordnet.princeton.edu/`

[10] `https://babelnet.org/` (Bevilacqua et al., 2021)

[11] `https://iate.europa.eu/home`

[12] `https://www.collinsdictionary.com/dictionary/english`

[13] `https://www.oxfordlearnersdictionaries.com`

[14] `https://www.wiktionary.org/`

[15] `https://eur-lex.europa.eu/browse/eurovoc.html?locale=en`

[16] `https://vocabularies.unesco.org/browser/thesaurus/en/`

[17] `https://agrovoc.fao.org/browse/agrovoc/en/`

Song, and Yang (2018) utilized the Thesaurus for Geographic Senses in conjunction with computational lexicons such as WordNet and HowNet (Dong and Dong, 2003) to compute semantic similarity between terms from diverse geographic domain terminologies. These examples highlight the versatility of thesauri in facilitating semantic integration across terminologies. Although thesauri are effective tools, they present several limitations: despite often including thousands of terms, they cannot guarantee comprehensive coverage, particularly in domains with emerging or undocumented terminologies, where the automatic and continuous integration of new terms remains a significant challenge.

5. Textual sources (raw text)-based: Through text mining approaches, textual resources specialized in a specific domain can be leveraged to enrich the semantic information of terms and facilitate their disambiguation. TermCork is a corpus-based system, designed to extract terms along with their definitions to link linguistic resources. This process involves gathering examples of usage, where terms and their corresponding equivalents can be observed in context, intending to assist linguists in the creation of glossaries or dictionaries. Similarly, this technique can be applied to generate lexical-semantic contexts for terms to be linked (Ramos, 2020). The accuracy of these mappings is highly dependent on corpus quality, meaning that results are limited by the size, specialization, and accuracy of the available textual resources, which may not be sufficient in specialized domains.

6. Sense inventories-based: Disambiguation can be approached in a traditional way, where each term is labeled with the meaning that corresponds to it based on the context in which it appears. Sense inventories serve as key tools to facilitate the identification and assignment of senses during the terminology mapping process. These inventories are predefined lexical databases that compile all possible meanings of words in one or more languages (Bevilacqua et al., 2021). Efforts have been made to develop specialized sense inventories that address the unique needs of particular domains. Examples of such inventories are the ones developed by Grossman et al. (2018) and Dong and Wang (2021). However, specialized sense inventories continue to exhibit some limitations. In the work by Bevilacqua, Maru, and Navigli (2020) several challenges associated with the use of sense inventories to address lexical-semantic ambiguity are identified: they are generally static and fail to reflect the dynamic changes in language use and lack consistent coverage across languages.

7. Domain inventories-based: Current research on domain-specific disambiguation systems often leverages domain label inventories and algorithms that repurpose large existing lexical resources to analyze specialized literature (Chopard, Corcoran, and Spasić, 2024). Domain inventories refer to collections of domain labels assigned to terms, many of which originate from computational lexicons such as WordNet Domains[18] or BabelDomains[19]. A notable example is the Coarse Sense Inventory (CSI) (Kikuchi et al., 2024), a resource developed to reduce the granularity of WordNet by consolidating word senses into broader categories. Despite its strengths, the domain information derived from WordNet in CSI remains too general to handle highly specialized terminologies, and its coverage is restricted to English.

4.4 Hybrid methods

According to World Health Organization (2021), hybrid methods are defined as the combination of two or more approaches that were presented in the previous sections. This approach is widely employed in various disambiguation tasks. All methods that incorporate manual evaluation as part of their process are considered hybrid methods, as they combine automated techniques with human oversight. Barba, Pasini, and Navigli (2021) argues that hybrid methods produce the best results in tasks that require a disambiguation process. A notable example of such methods is the MARIE system developed by Kim et al. (2020), which integrates string-matching techniques with contextual embedding approaches. This system links terms that share the same concept regardless of their domain while effectively disambiguating their meanings. Several other examples combine methods according to their available data (Ballatore, Bertolotto, and Wilson, 2014; Bulla et al., 2022; Stroganov et al., 2022).

[18] `https://wndomains.fbk.eu/`
[19] `http://lcl.uniroma1.it/babeldomains/`

5 LLM-based methods

LLMs have emerged as promising tools for terminology mapping, as they can generate rich contextual information and identify different senses of terms, enhancing disambiguation. However, despite their potential, we have not identified any studies specifically dedicated to terminology mapping, terminological variation, or the associated task of disambiguation. Therefore, we present the following studies, which are closely related and may serve as a bridge for establishing future research directions towards LLM-based terminology mapping.

Many challenges address the need of managing specialized terms for *machine translation* tasks, such as Moslem et al. (2023), that propose the use of LLMs to generate high-quality bilingual specialized synthetic terminologies, enabling the representation of specific senses of terms in such terminologies. A similar approach is the work of Bogoychev and Chen (2023), that suggests the use of LLMs to refine translations by generating domain-specific terminology through prompting strategies. Jhirad, Marrese-Taylor, and Matsuo (2023) further demonstrates that LLMs possess a deep understanding of specialized terminology across various domains, particularly in their ability to generate precise definitions by constructing a large dataset and designing a series of evaluation tasks in the financial domain. Additionally, Joachimiak et al. (2024) shows that LLMs can effectively map concepts and terms from specialized fields, such as the artificial intelligence domain, making them valuable tools for the development of new terminological resources due to their deep understanding of domain-specific knowledge. Another relevant study is Fan et al. (2024), that proposes a novel framework for linking mentions of clinical terms in textual sources with a medical termbase. By using LLMs, they address the problem of terminological variation, particularly in the normalization of disease diagnoses terminology.

6 Discussion and conclusions

In this survey, we have examined various approaches that address ambiguity and term variation to integrate heterogeneous terminological resources. Based on our analysis of the literature, we conclude that external resources can improve the disambiguation process by enriching the semantic information within terminologies. Additionally, hybrid approaches that combine lexical and semantic techniques are particularly effective in handling complex domain-specific terminologies (Q2).

Ambiguity and term variation stem from the various linguistic phenomena described in Section 2.1. These phenomena, together with inconsistencies in how terms are structured across different resources, represent the main challenges in achieving semantic integration of terminologies. Particularly, these challenges vary across domains of specialization. While the biomedical domain benefits from authoritative resources like UMLS and SNOMED-CT, other fields lack systematic identifiers and comprehensive terminological databases. As shown in Table 1, from the 21 studies analyzed, 12 focus on biomedical applications, reflecting a research bias towards resource-rich domains. Moreover, the structured formats used in biomedical terminologies are highly specialized and cannot be easily extrapolated to other fields, where terminological representation differs significantly. In non-biomedical domains, the absence of standardized resources complicates semantic interoperability, making it difficult to establish clear correspondences between terms (Q1).

Research on terminology mapping will require the development of more robust approaches that integrate innovative techniques with existing methodologies to improve accuracy and interoperability across diverse knowledge areas. As exemplified in Section 5, LLMs show great potential for this task by generating contextualized representations and identifying term senses. Given the persistent challenges in resolving ambiguity and addressing resource limitations, future research should explore LLMs to improve scalability and adaptability in terminology mapping (Q3).

Acknowledgements

This work has received funding from the INESData project (Infrastructure to Investigate Data Spaces in Distributed Environments at UPM), a project funded under the UNICO I+D CLOUD call by the Ministry for Digital Transformation and the Civil Service, within the framework of the recovery plan PRTR financed by the European Union (NextGenerationEU).

References

Alkhamaiseh, K. and S. ALShagarin. 2014. A survey of string matching algorithms. *International Journal of Engineering Research and Applications*, 4:144–156, 08.

Amin, A., M. Hildebrand, J. van Ossenbruggen, and L. Hardman. 2010. Designing a thesaurus-based comparison search interface for linked cultural heritage sources. In *Proceedings of the 15th international conference on Intelligent user interfaces*, pages 249–258.

Ballatore, A., M. Bertolotto, and Wilson. 2014. Linking geographic vocabularies through wordnet. *Annals of GIS*, 20, 04.

Barba, E., T. Pasini, and R. Navigli. 2021. ESC: Redesigning WSD with extractive sense comprehension. In K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tur, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, and Y. Zhou, editors, *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4661–4672, Online, June. Association for Computational Linguistics.

Baylor, E., E. Ploeger, and J. Bjerva. 2023. The past, present, and future of typological databases in NLP. In H. Bouamor, J. Pino, and K. Bali, editors, *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 1163–1169, Singapore, December. Association for Computational Linguistics.

Ben Aouicha, M., M. A. Hadj Taieb, and A. Ben Hamadou. 2016. Taxonomy-based information content and wordnet-wiktionary-wikipedia glosses for semantic relatedness. *Applied Intelligence*, 45, 09.

Besançon, R., G. de Chalendar, O. Ferret, F. Gara, O. Mesnard, M. Laïb, and N. Semmar. 2010. LIMA : A multilingual framework for linguistic analysis and linguistic resources development and evaluation. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis, M. Rosner, and D. Tapias, editors, *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, Valletta, Malta, May. European Language Resources Association (ELRA).

Bevilacqua, M., M. Maru, and R. Navigli. 2020. Generationary or "how we went beyond word sense inventories and learned to gloss". In B. Webber, T. Cohn, Y. He, and Y. Liu, editors, *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7207–7221, Online, November. Association for Computational Linguistics.

Bevilacqua, M., T. Pasini, A. Raganato, and R. Navigli. 2021. Recent trends in word sense disambiguation: A survey. In Z.-H. Zhou, editor, *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, pages 4330–4338. International Joint Conferences on Artificial Intelligence Organization, 8. Survey Track.

Bizer, C., T. Heath, and T. Berners-Lee. 2009. Linked data: The story so far. *International Journal on Semantic Web and Information Systems*, 5:1–22, 07.

Bogoychev, N. and P. Chen. 2023. Terminology-aware translation with constrained decoding and large language model prompting.

Bowker, L. 1997. 1.4.4 multidimensional classification of concepts and terms.

Bulla, L., M. C. Frangipane, M. L. Mancinelli, L. Marinucci, M. Mongiovì, M. Porena, V. Presutti, and C. Veninata. 2022. Developing and aligning a detailed controlled vocabulary for artwork. In *Symposium on Advances in Databases and Information Systems*.

Cabré, M. T. 1993. *La terminología: teoría, metodología, aplicaciones*. Editorial Antártida/Empúries.

Chen, Z., J. Song, and Y. Yang. 2018. An approach to measuring semantic relatedness of geographic terminologies using a thesaurus and lexical database sources. *ISPRS International Journal of Geo-Information*, 7(3).

Chopard, D., P. Corcoran, and I. Spasić. 2024. Word sense disambiguation of acronyms in clinical narratives. *Frontiers in Digital Health*, 6.

Clavel-Merrin, G. 2004. Macs (multilingual access to subjects): A virtual authority file across languages. *Cataloging & Classification Quarterly*, 39:323–330, 11.

Dechandon, D., A. Gerencsér, and M. R. Ruiz. 2020. Terminology: towards a systematic integration of semantics and metadata. *Translating and the Computer*, 41:10–26.

Doerr, M. 2001. Semantic problems of thesaurus mapping. *J. Digit. Inf.*, 1, 01.

Dong, Q. and Y. Wang. 2021. Enhancing medical word sense inventories using word sense induction: A preliminary study. In V. Gadepally, T. Mattson, M. Stonebraker, T. Kraska, F. Wang, G. Luo, J. Kong, and A. Dubovitskaya, editors, *Heterogeneous Data Management, Polystores, and Analytics for Healthcare*, pages 151–167, Cham. Springer International Publishing.

Dong, Z. and Q. Dong. 2003. Hownet - a hybrid language and knowledge resource. In *International Conference on Natural Language Processing and Knowledge Engineering, 2003. Proceedings. 2003*, pages 820–824.

Dybå, T., T. Dingsøyr, and G. Hanssen. 2007. Applying systematic reviews to diverse study types: An experience report. pages 225 – 234, 10.

Euzenat, J. and P. Shvaiko. 2013. *Ontology Matching*. Springer Berlin Heidelberg.

Fan, Y., Y. Zhu, K. Xue, J. Liu, and T. Ruan. 2024. RRNorm: A novel framework for Chinese disease diagnoses normalization via LLM-driven terminology component recognition and reconstruction. In L.-W. Ku, A. Martins, and V. Srikumar, editors, *Findings of the Association for Computational Linguistics: ACL 2024*, pages 9162–9175, Bangkok, Thailand, August. Association for Computational Linguistics.

Freixa, J. and S. Fernandez-Silva. 2017. Chapter 7. terminological variation and the unsaturability of concepts.

Freixa Aymerich, J. 2003. Variació terminològica: anàlisi de la variació denominativa en textos de diferent grau d'especialització de l'àrea de medi ambient, la. 03.

Fung, K. W., O. Bodenreider, A. Aronson, W. Hole, and S. Srinivasan. 2007. Combining lexical and semantic methods of inter-terminology mapping using the umls. *Studies in health technology and informatics*, 129:605–9, 02.

Fung, K. W. and J. Xu. 2015. An exploration of the properties of the core problem list subset and how it facilitates the implementation of snomed ct. *Journal of the American Medical Informatics Association*, 22(3):649–658, 02.

Grossman, L. V., E. G. Mitchell, G. Hripcsak, C. Weng, and D. K. Vawdrey. 2018. A method for harmonization of clinical abbreviation and acronym sense inventories. *Journal of Biomedical Informatics*, 88:62–69.

Health informatics. ISO/TC 215. 2014. ISO/TR 12300:2014. Technical Report. Health informatics — Principles of mapping between terminological systems. Technical report, ISO standards, November. Accessed: 2024-12-05.

Hertling, S. and H. Paulheim. 2023. Olala: Ontology matching with large language models. In *Proceedings of the 12th Knowledge Capture Conference 2023*, K-CAP '23, page 131–139. ACM, December.

Jhirad, J., E. Marrese-Taylor, and Y. Matsuo. 2023. Evaluating large language models' understanding of financial terminology via definition modeling. In D. Li, R. Mahendra, Z. P. Tang, H. Jang, Y. Murawaki, and D. F. Wong, editors, *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics: Student Research Workshop*, pages 93–100, Nusa Dua, Bali, November. Association for Computational Linguistics.

Joachimiak, M. P., M. A. Miller, J. H. Caufield, R. Ly, N. L. Harris, A. Tritt, C. J. Mungall, and K. E. Bouchard. 2024. The artificial intelligence ontology: Llm-assisted construction of ai concept hierarchies.

Kang, B., J. Yoon, H. Y. Kim, S. J. Jo, Y. Lee, and H. J. Kam. 2021.

Deep-learning-based automated terminology mapping in omop-cdm. *Journal of the American Medical Informatics Association*, 28(7):1489–1496, 05.

Kikuchi, M., M. Ono, T. Soga, T. Tanabe, and T. Ozono. 2024. Coarse-grained sense inventories based on semantic matching between english dictionaries.

Kilgarriff, A. and C. Yallop. 2000. What's in a thesaurus? In *International Conference on Language Resources and Evaluation*.

Kim, H. K., S. Choi, Y. Bae, J. Choi, H. Kwon, C. Lee, H.-Y. Lee, and T. Ko. 2020. Marie: A context-aware term mapping with string matching and embedding vectors. *Applied Sciences*, 10:7831, 11.

Kolitsas, N., O.-E. Ganea, and T. Hofmann. 2018. End-to-end neural entity linking. In A. Korhonen and I. Titov, editors, *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pages 519–529, Brussels, Belgium, October. Association for Computational Linguistics.

León-Araúz, P. 2017. Chapter 9. term and concept variation in specialized knowledge dynamics.

Mate, S., S. A. Seuchter, K. Ehrenberg, N. Deppenwiese, J. Zierk, H.-U. Prokosch, D. Kraska, and L. A. Kapsner. 2021. A multi-user terminology mapping toolbox. In *German Medical Data Sciences: Bringing Data to Life*. IOS Press, pages 217–223.

McCulloch, E., A. Shiri, and D. Nicholson. 2005. Challenges and issues in terminology mapping: A digital library perspective. *The Electronic Library*, 23, 12.

McInnes, B. T. and T. Pedersen. 2013. Evaluating measures of semantic similarity and relatedness to disambiguate terms in biomedical text. *Journal of biomedical informatics*, 46 6:1116–24.

Moslem, Y., G. Romani, M. Molaei, R. Haque, J. D. Kelleher, and A. Way. 2023. Domain terminology integration into machine translation: Leveraging large language models.

National Health Service. n.d. Cross-mapping. Online. [cited 2024-12-05].

Navigli, R. 2009. Word sense disambiguation: A survey. *ACM Comput. Surv.*, 41, 02.

Navigli, R. and P. Velardi. 2010. Learning word-class lattices for definition and hypernym extraction. In *Annual Meeting of the Association for Computational Linguistics*.

Nikiema, J. N., V. Jouhet, and F. Mougin. 2017. Integrating cancer diagnosis terminologies based on logical definitions of snomed ct concepts. *Journal of biomedical informatics*, 74:46–58.

Portisch, J., M. Hladik, and H. Paulheim. 2019. Wiktionary matcher. In *OM@ISWC*.

Qamar, R., A. L. Rector, Kilburn, and Building. 2006. Semantic mapping of clinical model data to biomedical terminologies to facilitate data interoperability.

Ramos, M. 2020. *Knowledge Organisation and Terminology: Application to Cork*. Ph.D. thesis, 11.

Schmidt, D., A. D. Bosco, C. Trojahn, R. Vieira, and P. Quaresma. 2020. Aligning iate criminal terminology to sumo. In *International Conference on Computational Processing of the Portuguese Language*.

Stroganov, O., A. Fedarovich, E. Wong, Y. Skovpen, E. Pakhomova, I. Grishagin, D. Fedarovich, T. Khasanova, D. Merberg, S. Szalma, and J. Bryant. 2022. Mapping of uk biobank clinical codes: Challenges and possible solutions. *PLOS ONE*, 17.

Vizine-Goetz, D., C. Hickey, A. Houghton, and R. Thompson. 2006. Vocabulary mapping for terminology services. *J. Digit. Inf.*, 4.

Volz, J., C. Bizer, M. Gaedke, and G. Kobilarov. 2009. Silk - a link discovery framework for the web of data. In *LDOW*.

Vrandečić, D. and M. Krötzsch. 2014. Wikidata. *Communications of the ACM*, 57:78 – 85.

Wang, Y., J. Patrick, G. Miller, and J. Gordon. 2012. Linguistic mapping of terminologies to snomed ct. 04.

World Health Organization. 2021. Who-fic classifications and terminology mapping: Principles and best practice. https://cdn.who.int/media/docs/default-source/classification/who-fic-network/whofic_terminology_mapping_guide.pdf?sfvrsn=2cae387c_7&download=true. Accessed November 25, 2024.

Zeng, M. and L. Chan. 2004. Trends and issues in establishing interoperability among knowledge organization systems. *JASIST*, 55:377–395, 03.