REVIEW ARTICLE

A survey of knowledge organization systems of research fields: Resources and challenges

Angelo Salatino[1], Tanay Aggarwal[1], Andrea Mannocci[2], Francesco Osborne[1,3], and Enrico Motta[1]

[1]Knowledge Media Institute, The Open University, Milton Keynes, UK
[2]Istituto di Scienza e Tecnologie dell'Informazione "Alessandro Faedo," Consiglio Nazionale delle Ricerche, Pisa, Italy
[3]Department of Business and Law, University of Milano-Bicocca, Milan, Italy

Citation: Salatino, A., Aggarwal, T., Mannocci, A., Osborne, F., & Motta, E. (2025). A survey of knowledge organization systems of research fields: Resources and challenges. *Quantitative Science Studies*, *6*, 567–610. https://doi.org/10.1162/qss_a_00363

DOI:
https://doi.org/10.1162/qss_a_00363

Peer Review:
https://www.webofscience.com/api/gateway/wos/peer-review/10.1162/qss_a_00363

Received: 6 September 2024
Accepted: 10 February 2025

Corresponding Author:
Angelo Salatino
angelo.salatino@open.ac.uk

Handling Editor:
Vincent Larivière

Keywords: controlled vocabularies, digital libraries, knowledge organization systems, scholarly knowledge, scholarly ontologies, taxonomies

ABSTRACT

Knowledge organization systems (KOSs), such as term lists, thesauri, taxonomies, and ontologies, play a fundamental role in categorizing, managing, and retrieving information. In the academic domain, KOSs are often adopted for representing research areas and their relationships, primarily aiming to classify research articles, academic courses, patents, books, scientific venues, domain experts, grants, software, experiment materials, and several other relevant products and agents. These structured representations of research areas, widely embraced by many academic fields, have proven effective in empowering AI-based systems to enhance the retrievability of relevant documents, enable advanced analytic solutions to quantify the impact of academic research, and analyze and forecast research dynamics. We aim to present a comprehensive survey of the current KOS for academic disciplines. We analyzed and compared 45 KOSs according to five main dimensions: scope, structure, curation, usage, and links to other KOSs. Our results reveal a very heterogeneous scenario in terms of scope, scale, quality, and usage, highlighting the need for more integrated solutions for representing research knowledge across academic fields. We conclude by discussing the main challenges and the most promising future directions.

1. INTRODUCTION

Knowledge organization systems (KOSs), such as term lists, thesauri, taxonomies, and ontologies, play a fundamental role in categorizing, managing, and retrieving information (Mazzocchi, 2018). Specifically, they "model the underlying semantic structure of a domain and provide semantics, navigation, and translation through labels, definitions, typing, relationships, and properties for concepts" (Zeng, 2008).

In the academic domain, KOSs are often adopted for representing research areas and their relationships (Sugimoto & Larivière, 2018), with the primary aim of classifying research articles, books, courses, patents, grants, software, experiment materials, scientific venues, domain experts, organizations, and several other relevant items and agents. These structured representation of knowledge have been adopted by most academic fields and proved very effective in improving the retrievability of relevant documents (Newman, Noh et al., 2010; Salatino, Osborne et al., 2019); enabling advanced analytic solutions to quantify the impact of research (Ding, Rousseau, & Wolfram, 2014; Qiu, Zhao et al., 2017; Salatino, Angioni et al., 2023;

The MIT Press

Sugimoto & Larivière, 2018); and understanding and forecasting research dynamics (Qiu et al., 2017; Scharnhorst, Börner, & van den Besselaar, 2012).

More recently, these KOSs have become even more instrumental given the fast-growing number of publications, the rise of Open Science and Open Access articles, the thriving role of interdisciplinary research, and the emergence of vast online repositories of articles, academic courses, and other academic materials (Auer, Kovtun et al., 2018). This transformation poses new opportunities but also new challenges. For example, in the recent COVID-19 pandemic, there was a lot of discussion on how the scientific community was "drowning in COVID-19 papers" and had to resort to new tools based on robust representations of research concepts (Brainard, 2020). To address these issues, KOSs have been increasingly incorporated into various AI systems to assist researchers in navigating literature (Dai, You et al., 2020) and semiautomating systematic reviews (Bolaños, Salatino et al., 2024). Despite the emergence of new AI systems based on large language models (LLMs) in the last 2 years, structured and machine-readable representations of domain knowledge continue to be invaluable, as they aid in formulating precise queries to identify relevant publications, reduce hallucinations, and enhance interpretability (Ayala & Bechard, 2024; Gnoli, Golub et al., 2024).

KOSs of research areas are very heterogeneous in terms of scope, scale, quality, and usage. Some fields (e.g., "Biomedical") are well covered by a variety of KOSs (e.g., *MeSH, UMLS, NLM*) that are used to categorize research products and are routinely adopted by libraries, online repositories, researchers, and organizations. Other fields (e.g., "Mathematics") prevalently rely on one widely accepted KOS (e.g., *Mathematics Subject Classification*). A few research areas (e.g., "Geography," "History," "Material Science," "Political Science," and "Sociology") do not even have their own specific KOSs. Several large KOSs cover multiple academic disciplines, but often with coarse-grained representation that is not sufficient for the needs of the specific fields. To the best of our knowledge, we still lack a systematic and in-depth analysis of these knowledge organization systems and their characteristics.

The objective of this paper is to present a comprehensive survey of KOSs for academic fields. We defined formal inclusion and exclusion criteria that led to the identification of 45 candidates. We analyzed them according to five main aspects. The first is the **scope** of a KOS in terms of its coverage of academic disciplines. The second aspect is the **structure**, which includes features such as the number of concepts, the maximum depth, the type of hierarchy, and the presence of synonyms. The third aspect is **curation**, which includes information about the formats, the license, the frequency of updates, the procedure used for its generation, and the languages in which the KOS is available. The fourth aspect deals with the **links to other KOSs**, which allow users and tools to interconnect and adopt multiple KOSs for a richer characterization of research areas. The final aspect regards their **usage** in digital libraries, repositories, and research communities.

We conclude the paper by discussing the main challenges and opportunities in this field, highlighting the most promising future directions. We also analyze how current solutions could be integrated and interlinked to produce a more comprehensive and granular representation of all academic disciplines.

In adherence to Open Science principles, we release the table that describes all the identified KOSs according to the 15 features of the Open Research Knowledge Graph[1], as well as the code[2] we developed for processing them.

[1] Full table describing the analyzed KOSs according to the 15 features: https://doi.org/10.48366/R732033.
[2] The code for processing the analyzed KOSs is available on GitHub: https://github.com/angelosalatino/kos-rf.

The rest of this manuscript is organized as follows. Section 2 introduces KOSs and discusses their applications. Section 3 describes the methodology that we employed to identify the 45 KOSs and introduces the full set of features used for the analysis. Section 4 presents the results of our analysis, offering a thorough, feature-by-feature assessment of the current landscape, and illustrating significant patterns. Section 5 discusses the ongoing challenges and possible future directions. Section 6 outlines the threats to the validity of our analysis. Finally, Section 7 concludes the paper by summarizing the contributions and the main findings.

2. BACKGROUND

KOSs play a crucial role in research by providing structured frameworks for organizing complex information, allowing researchers to establish clear categories, discern relationships, and navigate large data sets with increased efficiency. To underscore their importance, we will draw upon the literature to describe their usage in the research domain, specifically focusing on two main angles: digital libraries (Section 2.1) and information science (Section 2.2).

2.1. Knowledge Organization Systems in Digital Libraries

Knowledge organization systems form the backbone of effective search and retrieval in digital libraries, providing a systematic means for categorizing and organizing knowledge, retrieving information, facilitating preservation, and ensuring interoperability (Hodge, 2000). Annotating research products with appropriate research concepts facilitates semantic searches, leading to more effective information retrieval.

In the literature, we can find different types of KOSs, such as taxonomies, glossaries, dictionaries, synonym rings, gazetteers, authority files, subject headings, thesauri, classification schemes, semantic networks, and ontologies (Hodge, 2000; Zeng, 2008). Zeng (2008) comprehensively emphasizes the interplay between the complexity of their structures and their expected functions. The complexity of their structure can range from simple "flat" structures (e.g., *pick lists, dictionaries*), to two-dimensional hierarchical structures (e.g., *taxonomies*), and finally, to multidimensional structures, creating networks according to diverse semantic types and relationship (e.g., *ontologies*). Generally, KOSs with higher structural complexity exhibit greater capacity to suit various functions, including disambiguation of terms, management of synonyms or equivalent terms, establishment of semantic relationships, particularly hierarchical and associative links, and representation of both conceptual relationships and attributes within knowledge models. We refer the interested reader to Hodge's book (Hodge, 2000) and Zeng (2008) for additional details on the various types of KOSs.

Because KOSs are means for organizing information, they are at the heart of every digital library (Hodge, 2000). Indeed, well-known publishers like Elsevier, Springer Nature, the Institute of Electrical and Electronics Engineers (IEEE), and the Association for Computing Machinery (ACM) have developed their own systems to provide the full text of documents linked to bibliographic records. Major bibliographic databases like Web of Science, Scopus, Dimensions, and OpenAlex also employ KOSs to organize their vast collections of bibliographic records.

The annotation of documents based on the concepts within KOSs can be performed either manually or automatically. Manual annotation tasks are undertaken by human experts, typically experienced curators or editors, who leverage their domain knowledge to critically assess document content and assign the most pertinent concepts. In contrast, automatic annotation employs a range of computational tools that often incorporate advanced artificial intelligence techniques (Salatino et al., 2019). For instance, OpenAlex, which is a major bibliographic catalogue of scientific papers, employs a deep learning model that, based on research papers' title, abstract, citations, and journal name, can define the appropriate topics drawn from the *OpenAlex Topics* vocabulary (OpenAlex, 2024).

In addition to the automatic classification of content, KOSs also support additional tasks, such as augmented retrieval (Shiri, Revie, & Chowdhury, 2002), recommender systems (Cleverley & Burnett, 2015), integration and interoperability (Zeng & Mayr, 2019), and knowledge management and preservation (Chowdhury, 2010; Kopácsi, Hudak, & Ganguly, 2017). For augmented retrieval, KOSs support precision search and query expansion by allowing users to execute highly specific queries using controlled vocabulary and discover more relevant content. In particular, by leveraging the related terms or broader concepts within KOSs, users can either manually expand their searches through user interfaces (Shiri et al., 2002) or rely on the search engine to automatically expand their queries (Mu, Lu, & Ryu, 2014).

With regard to recommender systems, KOSs enable the development of applications that enhance content discovery by suggesting related content based on subject matter and providing personalized recommendations derived from user search patterns, fostering serendipitous discovery and richer user engagement (Cleverley & Burnett, 2015; Thanapalasingam, Osborne et al., 2018).

In the context of integration and interoperability, KOSs establish a framework for the semantic enrichment of data, facilitating seamless integration of research products across different digital libraries or repositories and consequently enhancing their interoperability (Zeng & Mayr, 2019).

Finally, KOSs can also contribute to the long-term preservation of information by ensuring it is organized logically and can be easily retrieved and understood in the future (Chowdhury, 2010). In this regard, Kopácsi et al. (2017) argue that adding standardized values as metadata, selecting them from predefined controlled vocabularies rather than guessing keywords, improves the long-term preservation of digital objects.

In conclusion, digital libraries often employ KOSs to improve document organization and provide a wide range of advanced features.

2.2. Knowledge Organization Systems in Information Science

The research community has utilized KOSs of research topics to enable and support a variety of tasks in this domain, such as analysis of the scientific landscape (Angioni, Salatino et al., 2022a; Reymond, 2020; Yang & Lee, 2018), trend analysis and forecasting (Salatino, Osborne, & Motta, 2018; Yan, 2014), analyze the composition of a research team (Salatino et al., 2023), and assessing impact (Sjögårde & Didegah, 2022). Here we outline a small sample of approaches employed for these tasks.

In the context of analyzing the scientific landscape, Angioni et al. (2022a) developed the AIDA Dashboard, a tool that facilitates the analysis of conferences and journals in Computer Science, providing valuable insights into main authors, organizations, and countries. The dashboard leverages the *Computer Science Ontology* (Salatino, Thanapalasingam et al., 2018) to provide a very granular representation of the venues' research topics, as well as to rank venues within topics using various metrics. Further contributions include Reymond (2020), who examined patents in the Humanities using the *UNESCO Thesaurus*. Their findings provided useful insights into potential research questions and unexplored research avenues. Furthermore, the work of Yang and Lee (2018) introduced a tool that assists users in analyzing research trends by expanding initial queries using *MeSH* terms, facilitating a more comprehensive exploration of the research landscape.

For research trends analysis, Ilgisonis, Pyatnitskiy et al. (2022) performed a systematic retrospective analysis of the frequencies of *MeSH* concepts across 12 years. Their analysis revealed potential shifts in scientific priorities, and they employed the same patterns to predict emerging trends within a 5 year timeframe. In addition, Ovalle-Perandones, Gorraiz et al. (2013) studied whether the European Framework Programmes shaped the scientific output in "nanotechnology" of its member states. Their study compared this output to global trends as well as patterns of international collaboration. The authors relied on the representations of "nanotechnology" within *EuroVoc, MeSH*, and three additional KOSs to construct a refined search query for retrieving relevant papers from Web of Science.

Within the analysis of research team composition, Salatino et al. (2023) investigated how the diversity of expertise of a research team can influence their scientific impact. In this experiment, research topics from the *Computer Science Ontology* (Salatino, Thanapalasingam et al., 2018) were employed to model the researcher's expertise. Specifically, they characterize the expertise of an author at the time of collaboration as the distribution of research topics of their paper over the preceding 5 years. Additionally, Kang, Li, and Coppel (2015) mapped researchers' expertise according to *MeSH* terms and developed a matching algorithm to find potential interdisciplinary collaborators.

Regarding citation impact, Sjögårde and Didegah (2022) performed an analysis of the correlation between topic growth and article citation. Their methodology leverages an in-house, automatically generated KOS (Sjögårde & Ahlgren, 2020) to systematically categorize topics and disciplines within their data set. Their topic-based analysis yields compelling insights that could shape future research policy decisions. Moreover, Chatzopoulos, Vergoulis et al. (2020) introduced a method for estimating the impact of recently published papers, based on the premise that similar papers often experience comparable popularity trajectories. In this context, they leveraged citation networks and metadata, including the *Computer Science Ontology*, to assess similarity.

These analyses and approaches underscore the value and importance of KOSs in research, as they provide the essential structure and organization needed for developing various downstream applications (Salatino, Osborne, & Motta, 2020).

3. SURVEY METHODOLOGY

In this section, we describe the process we followed to identify and analyze KOSs of academic disciplines. We first define the types of KOSs that are objects of this analysis (Section 3.1) and describe the inclusion and exclusion criteria adopted in the survey (Section 3.2). We then describe the strategy we adopted to find the KOSs (Section 3.3) and discuss the set of features that we will use for describing and comparing them (Section 3.4).

To ensure clarity and consistency, we adopt the following typographical conventions: KOSs are indicated in italics (e.g., *Medical Subject Headings*) while research concepts (topics and disciplines) are enclosed in double quotation marks (e.g., "Medicine"). These conventions are designed to help readers easily distinguish between different types of entities.

3.1. Concepts

In this survey, we focus on KOSs of academic fields. KOSs are commonly used to categorize items into specific classes. In this context, classes correspond to research topics within various disciplines. The items categorized can include a wide range of artefacts, such as documents (e.g., research papers, patents, project reports), videos (e.g., university courses), data sets, software, projects, and more. As proposed in Salatino (2019), we define a **research topic** as the subject of study or issue that is of interest to the academic community, and it is explicitly addressed by research papers. Examples of research topics are "Acoustics"[3] in the *PhySH*[4] taxonomy, "Hydrocodone"[5] in *MeSH*[6], or "Web Ontology Language (OWL)"[7] in *ACM CCS*[8]. On the other hand, we define a **research field** or **discipline** as a broad area of knowledge within academia that consists of several research topics. For instance, "Mathematics" and "Medicine" are disciplines encompassing various and more specific research topics such as "algebra," "calculus," "oncology," and "cardiology."

KOSs in this space primarily fall into four possible **types of system**: term lists, hierarchical taxonomies, thesauri, and ontologies. In the following, we describe each one of them.

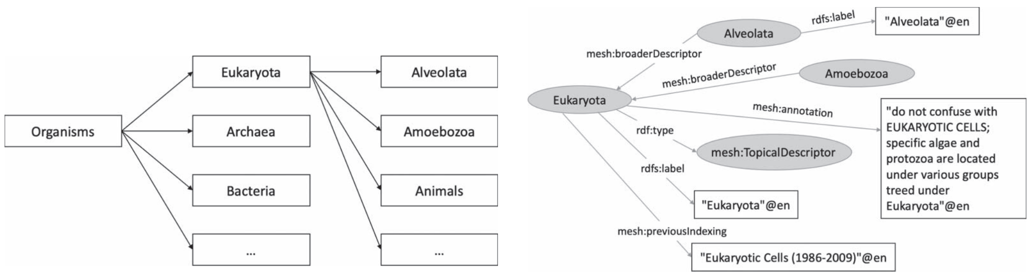
A **term list** is a flat list of subject headings or descriptors that support the organization of a collection of documents (Hedden, 2010; Zaharee, 2013). The ANSI/NISO Z39.19-2005 standard refers to it also as "pick list" (ANSI/NISO Z39.19-2005(R2010); National Information Standards Organization, 2010), whereas within the ISO 25964-2, it is referred to as "terminology" (BS ISO 25964-2:2013; International Organization for Standardization, 2013). The most important distinction from other types of KOS is that term lists do not define any relationships between the subjects.

A **hierarchical taxonomy** (from here on just *taxonomy*) organizes the classes in a hierarchical structure with parent-child relationships (Rasch, 1987). In practical terms, as shown in Figure 1(a), a taxonomy is typically organized in a tree structure, with a root node at the top that unfolds in several subbranches. An important characteristic is that all items classified according to a class can also be considered under all its superclasses. For instance, in Figure 1(a), all "Bacteria" are also "Organisms."

A **thesaurus** can be seen as an extension of the taxonomy. Subjects are organized in a hierarchical structure and they can also be described according to additional properties, such as a description, related terms, and synonyms (ANSI/NISO Z39.19-2005(R2010); National Information Standards Organization, 2010).

Finally, an **ontology**, from the "Information Technology" perspective, is a formal and explicit description of knowledge within a domain, categorizing things according to their essential or relevant qualities (Gruber, 1993). Ontologies conceptualize a subject area according to an abstract model of how domain experts rationalize knowledge in the domain. In practice, ontologies consist of a set of concepts, objects, and other entities and the relationships between them (Genesereth & Nilsson, 2012). Compared to thesauri, ontologies are more expressive because they can also express axioms and restrictions, and so provide local constraints on properties. As an example, Figure 1(b) shows a portion of the *MeSH* ontology. In particular, the two `mesh:broaderDescriptor` relationships define the superclass of both "Alveolata" and "Amoebozoa" which is "Eukaryota." Next, `rdfs:label` provides a human-readable version of a given class. Moreover, the `mesh:previousIndexing` points to the

[3] Acoustics: https://physh.aps.org/concepts/40a5bd01-6544-4502-8321-458c33878bf3.
[4] PhySH: https://physh.aps.org.
[5] Hydrocodone: https://meshb.nlm.nih.gov/record/ui?ui=D006853.
[6] MeSH: https://meshb.nlm.nih.gov.
[7] Web Ontology Language (OWL): https://dl.acm.org/topic/ccs2012/10002951.10003260.10003309.10003315.10003316.
[8] ACM CCS: https://www.acm.org/publications/class-2012.



(a) Taxonomy **(b)** Ontology

Figure 1. A small portion of the Medical Subject Headings represented both through a taxonomy (a) and an ontology (b). The taxonomy consists solely of a set of terms connected by hierarchical relationships. The ontology is instead a rather complex structure as it contains entities (grey ovals) and literals (white boxes), connected through semantic relationships.

same element in a previous version of the KOS, and the `mesh:annotation` provides a human-readable description of the class.

In certain circumstances, as we will see from the analysis, the naming conventions are somewhat disorganized. Some KOSs are named after a particular category but, upon closer examination, actually represent a different category based on the definitions discussed above. For instance, the *UNESCO Thesaurus* and the *STW Thesaurus for Economics*, despite their names, have evolved away from traditional thesaurus structures, aligning with the expressiveness of ontologies (Martínez-González & Alvite-Díez, 2019).

3.2. Selection Criteria

In this section, we define standard inclusion (IC) and exclusion (EC) criteria for the survey.

We selected all KOSs which match the following inclusion criteria:

IC 1. They describe academic research topics as defined in Section 3.1;

IC 2. They cover at least one of the following 19 broad research fields: “Art,” “Biology,” “Business,” “Chemistry,” “Computer science,” “Economics,” “Engineering,” “Environmental science,” “Geography,” “Geology,” “History,” “Materials science,” “Mathematics,” “Medicine,” “Philosophy,” “Physics,” “Political science,” “Psychology,” and “Sociology”;

IC 3. They are adopted by the main bibliographic databases or regularly used by the scientific community.

For the **IC 2**, we adopted the top-level fields of the Microsoft’s Fields of Study (Sinha, Shen et al., 2015) due to their comprehensive coverage.

In contrast, we excluded KOSs that match the following criteria:

EC 1. Do not offer an English version;

EC 2. Are exclusively tailored to the content of a specific library and therefore are not adopted by a community of users.

We established the **EC 2** criterion because there is an abundance of KOSs created by specific libraries (e.g., the *Aarhus University Library Classification System*[9]) for internal needs, but not available or (re)used by the scientific community.

[9] Aarhus University Library Classification System: https://web.archive.org/web/20210721074627/https://library.au.dk/en/subject-areas/political-science/classification-system.

3.3. Methodology for the Retrieval of Relevant KOSs

Identifying all the relevant KOSs has proven to be a challenging task. A common approach, when pursuing a survey or a review of the literature, is to rely on a search engine (e.g., Scopus, Web of Science, Google Scholar) and construct a particular query that would return all research papers reporting the objects under review (Pranckutė, 2021). However, a good number of systems organizing research areas are not well described or documented in research papers. Therefore, relying on this typical approach would have produced limited results. To this end, we designed and performed the following systematic strategy.

Phase 1: We started by querying Google Scholar for potential candidates using the following query "*(controlled vocabulary OR taxonomy OR thesaurus OR ontology OR subject headings OR subject classification) AND (research OR science)*". This step identified seven KOSs.

Phase 2: We analyzed the websites of academic publishers, preprint archives, and academic search engines to identify the KOSs they use to organize their content. Specifically, we considered Scopus, Web of Science, Dimensions.ai, PubMed, the Springer Nature portal, ACM, IEEE, OpenAlex, ArXiv, and the last instance of Microsoft Academic Graph. This led to the identification of nine additional KOSs.

Phase 3: We adopted the Google search engine to identify additional KOSs, using the query "*("controlled vocabulary" OR "term list" OR "taxonomy" OR "thesaurus" OR "ontology" OR "subject headings" OR "subject classification")*" in combination with the 19 broad fields listed in **IC 2**. As a result, we identified 12 additional KOSs.

Phase 4: We contacted researchers working in the field of "Digital Libraries" and asked them whether they could point us toward any additional effort. They suggested two more KOSs.

Phase 5: We expanded the resulting KOSs by analyzing their links on Wikipedia. Specifically, we relied on the "See also" section, typically listing online databases and other related KOSs, which allowed us to discover five additional ones.

Phase 6: We retrieved and considered all the KOSs that are explicitly connected, or otherwise referenced into the KOSs identified at previous stages. As a result, we gathered eight additional KOSs.

Phase 7: Whenever we could not find at least one KOS within one of the 19 disciplines defined in **IC 2** (e.g., "History"), we reached out to professors in the missing disciplines, asking which KOS they employ, if any. Their response directed us toward two additional multifield KOSs.

3.4. Methodology for the Analysis of KOSs

We analyzed each KOS according to the five main aspects summarized in Figure 2: *Scope, Structure, Curation, External links*, and *Usage*.

3.4.1. Scope

The scope is the set of research fields covered by a KOS. Some KOSs focus on one specific field, such as *PhySH* for "Physics" or *Mathematics Subject Classification* for "Mathematics." These may also include elements from other complementary fields. This is the case of *PhilPapers Taxonomy*, which mainly focuses on "Philosophy" but also describes some concepts from other fields, including "Mathematics." Another category of KOSs covers multiple fields by

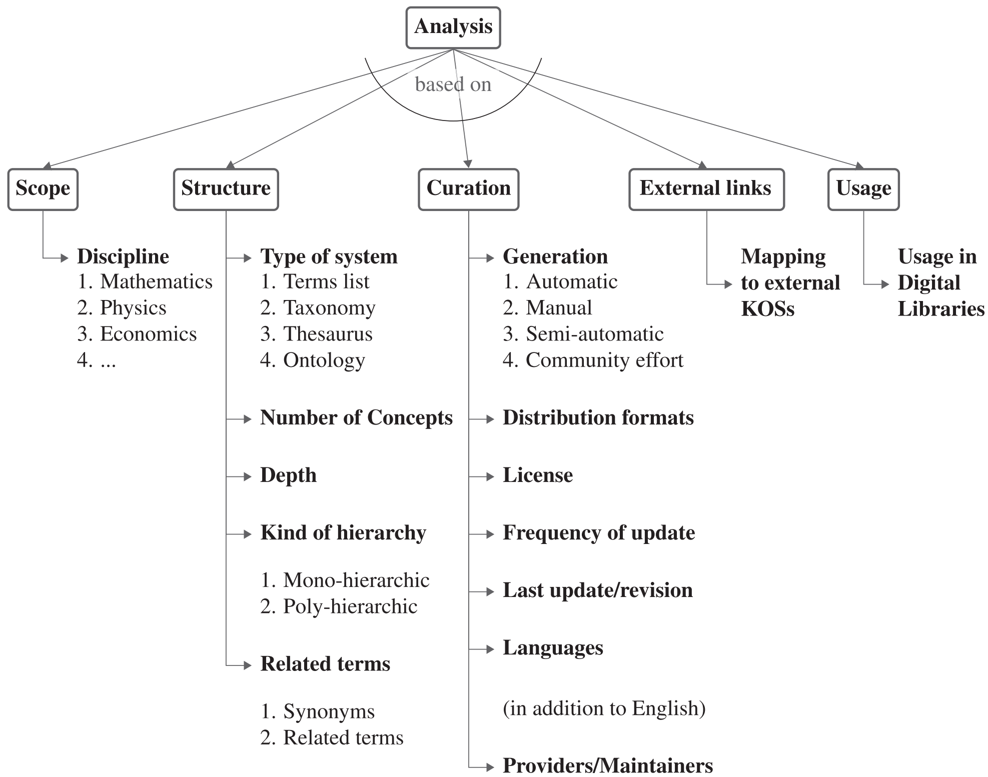


Figure 2. The aspects and features used for the analysis of KOSs.

design, typically because they aim to offer good coverage of an extended set of scientific or academic disciplines. A good example is the *UNESCO Thesaurus*, a well-known KOS that aims to cover all academic disciplines and is adopted by several libraries.

3.4.2. Structure

The structural characteristics of a KOS include many aspects based on its topology and the way subjects are arranged. First, we classified each KOS according to the four types defined in Section 3.1: term lists, taxonomies, thesauri, and ontologies. For example, the *Web of Science Categories* is a flat list of terms, and thus, we characterized it as a term list. The *Mathematics Subject Classification* and ANZSRC's *Fields of Research* are taxonomies because they arrange topics in a hierarchy. The *IEEE Thesaurus* is a thesaurus as it offers hierarchical information as well as synonyms. Finally, *TheSoz* is an ontology because, in addition to the hierarchical structure and synonyms, it adds axioms and constraints.

We then considered the number of concepts and the maximum depth of the hierarchical tree (which is one in the case of term lists). The depth was computed as the number of levels from the most generic concept (root) to the most specific concepts (leaves). Depth can be generally used as an indicator of specificity, as deep taxonomies tend to include a granular representation of very detailed (narrow) topics and may allow for more fine-grained content organization.

If the KOS is hierarchical, we also characterized it as either monohierarchical or poly-hierarchical according to the type of employed hierarchy. In monohierarchical KOSs, each concept is assigned to a single parent category. In contrast, poly-hierarchical KOS enable concepts to belong to multiple parent categories. As we will see, some KOSs attempt to organize their concepts within a strict monohierarchical structure (e.g., the *Mathematics Subject*

Classification), while others (e.g., *STW Thesaurus for Economics, Medical Subject Headings*) are organized in a poly-hierarchical structure.

We also considered if the KOS contains synonyms that can be used to refer to the same concept or related terms. For instance, the *Computer Science Ontology* uses "Ontology Matching" and "Ontology Mapping" as different labels for the same research topics. The *IEEE Thesaurus* offers instead a set of related terms that are not necessarily synonyms but are semantically similar. For instance, "4G mobile communication" is a related term of "5G mobile communication." The presence of different labels typically facilitates the classification of documents by both human annotators and automatic approaches.

We used Python scripts for automatically extracting the structural features from the KOSs published in standard formats (e.g., RDF). The scripts are publicly available as Jupyter notebooks on a GitHub repository[10]. For the knowledge organization systems that are available only as HTML or PDF, we manually extracted these features.

3.4.3. Curation

The curatorial aspect takes into consideration many features regarding the creation, update, and publication of the KOS. First, we considered how the KOS was generated according to four categories: manual, automatic, semiautomatic and community effort. The first and more common category includes all the KOSs that were crafted manually by domain experts, such as the *ACM Computing Classification Scheme*. The *automatic* category covers the KOSs that were generated automatically using NLP or Machine Learning approaches, such as the *Computer Science Ontology*. The *semiautomatic* category covers KOSs generated by blending domain experts and automatic tools. For instance, *OpenAlex Topics* uses a manual approach for the first three levels and an automatic one for the last level. The *community effort* category includes the KOSs that were crowdsourced with the support of the community, such as *Physical Subject Headings*.

We also considered the format used for publishing the KOSs. A good number of KOSs are published according to W3C standard formats such as Ontology Web Language (OWL), Terse RDF Triple Language (TTL), and N-Triples (NT) for expressing data in the Resource Description Framework (RDF)[11] data model, as well as JSON and CSV. An alternative is the OBO format[12], which is still a working draft standard for representing ontologies, based on the principles of OWL and predominantly used in the field of "Biology." Other KOSs are instead published in semi-structured formats, for instance, as a list of terms in HTML pages or PDF files. These solutions are typically not machine-interpretable and offer very limited support to automatic classifiers.

Next, we observed the distribution license. Some KOSs are open and released under Creative Commons licenses[13], while some others are copyrighted and may require a subscription fee. We also reported the frequency of updates and the date of the latest release, which is a good indication of whether the KOS is actively maintained. In addition, we considered whether the KOS is distributed in languages other than English. This is crucial for the interoperability of digital libraries across the different languages. Finally, we reported the current maintainers of those KOSs, as they are the first point of contact for interested readers. All of these curatorial features have been identified from the official websites and relevant literature.

[10] Jupyter notebooks for analyzing KOSs: https://github.com/angelosalatino/kos-rf.
[11] Resource Description Framework (RDF): https://www.w3.org/RDF.
[12] OBO Format: https://purl.obolibrary.org/obo/oboformat/spec.html.
[13] Creative Commons: https://creativecommons.org.

3.4.4. External links

Links to other KOSs allow the integration of different knowledge bases, potentially offering a more comprehensive representation of scientific disciplines. Therefore, we documented whether a KOS includes links to other KOSs, including those referencing general knowledge graphs, such as DBpedia[14] and Wikidata[15]. For example, the *STW Thesaurus for Economics* is mapped to Wikidata, DBpedia, and others. We discuss the crucial implications of this mechanism in Section 5.

3.4.5. Usage

The final aspect that we considered regards the presence of collections of documents or other artefacts that are annotated or organized according to the KOS. Their existence attests to the fact that a portion of the community actively uses that KOS. Furthermore, annotated data sets can also be used for training machine learning classifiers able to categorize new documents according to the original KOS (Kandimalla, Rohatgi et al., 2021; Salatino, Osborne et al., 2022). For instance, *Fields of Research* is currently adopted by institutional repositories in both Australia and New Zealand as well as Dimensions.ai[16] to index their publications metadata. Another example is *Physical Subject Headings* employed to classify the articles of the *Physical Review* journal[17].

4. ANALYSIS OF KNOWLEDGE ORGANIZATION SYSTEMS

This section analyzes the KOSs of research fields utilizing the 15 features reported in Figure 2. We identified 45 KOSs: 23 specializing in single disciplines and 22 covering multiple fields. Table 1 summarizes these KOSs according to some key characteristics (i.e., primary discipline, number of concepts, hierarchical depth, type of system, kind of hierarchy, and whether it contains related terms). An extended version of this table with all the analyzed features is available at https://doi.org/10.48366/R732033.

This section presents the outcomes of our analysis, structured according to the hierarchical structure illustrated in Figure 2. The subsections correspond to the five key aspects introduced in Section 3.4 (Scope, Structure, Curation, External links, and Usage) and are further divided into smaller subsections, each addressing a specific feature.

It is important to note that the results presented in this section are based on an analysis conducted up to April 2024. Due to the dynamic nature of some KOSs, minor discrepancies in values may have occurred since then; however, the overall insights and conclusions of this manuscript remain valid.

4.1. Scope

Out of the 45 identified KOSs, 22 cover multiple fields, while 23 focus on a single field of science. However, only 12 of the 19 fields under analysis (see IC 2, Section 3.2) are addressed by at least one single-field KOS, while the remaining ones rely exclusively on broader multi-field solutions.

Notably, six fields are covered by more than one KOS. In "Medicine," we identified four single-field KOSs: *Medical Subject Headings*, the *Unified Medical Language System*, the

[14] DBpedia: https://wiki.dbpedia.org.
[15] Wikidata: https://www.wikidata.org.
[16] Dimension.ai: https://app.dimensions.ai/discover/publication.
[17] *Physical Review* journal: https://journals.aps.org.

Table 1. An overview of the 45 knowledge organization systems, reporting their main discipline, number of concepts, depth, type of system (**Sy** = **O**ntology, **T**axonomy, ThesaUrus, and Term **L**ist), kind of hierarchy (**Hr** = **P**oly-hierarchical or **M**onohierarchical), and related terms (**RT** = **Y**es or **N**o)

Knowledge organization system	Main discipline	#Concepts	Depth	Sy	Hr	RT
Agrovoc Thesaurus	Agriculture	41,000	14	O	P	Y
Art and Architecture Thesaurus	Art & Architecture	58,000	13	O	P	Y
EDAM	Bioinformatics	264	7	O	P	Y
Open Biological and Biomedical Ontology	Biology	5.3 million	39	O	P	Y
Biomedical Ontologies from BioPortal	Biomedicine	13 million	–	O	P	Y
ChemOnt	Chemistry	4,825	11	O	M	Y
ACM Computing Classification Scheme	Computer Science	2,114	6	T	P	N
Computer Science Ontology	Computer Science	14,000	13	O	P	Y
Computer Science Subject Headings from Wikipedia	Computer Science	7,354	20	T	M	N
Journal of Economic Literature	Economics	859	3	T	M	N
STW Thesaurus for Economics	Economics	14,000	13	O	P	Y
IEEE Thesaurus	Engineering	12,000	3	U	P	Y
GeoRef Thesaurus	Geology	33,000	5	U	M	Y
U.S. Geological Survey Library Classification System	Geology	2,000	6	T	M	N
Mathematics Subject Classification	Mathematics	6,598	3	T	M	Y
Medical Subject Headings	Medicine	30,000	13	O	P	Y
National Library of Medicine classification	Medicine	4,761	4	T	M	N
Unified Medical Language System	Medicine	3.3 million	30	U	P	Y
PhilPapers Taxonomy	Philosophy	7,447	5	T	P	N
Physical Subject Headings	Physics	3,749	6	O	P	Y
Physics and Astronomy Classif. Scheme	Physics	8,260	5	T	M	N
PsycInfo and PsycTests Classification Systems	Psychology	188	3	T	M	N
TheSoz	Social Science	8,223	6	O	P	Y
All Science Journal Classification Codes	Multifield	334	2	T	M	N
ArXiv Subjects	Multifield	176	3	T	M	N
Dewey Decimal Classification	Multifield	60,000	14	T	P	Y
DFG Classification	Multifield	278	4	T	M	N
European Commission Taxonomy	Multifield	629	3	T	M	N

Table 1. (*continued*)

Knowledge organization system	Main discipline	#Concepts	Depth	Sy	Hr	RT
European Research Council Taxonomy	Multifield	431	3	T	M	N
EuroVoc	Multifield	7,439	6	O	P	Y
Fields of Research (ANZSRC)	Multifield	4,406	3	T	M	N
Fields of research and development (OECD)	Multifield	48	2	T	M	N
KNOWMAK	Multifield	72	3	O	P	Y
Library of Congress Class. (and Subj. Head.)	Multifield	467,000	29	O	P	Y
Microsoft's Fields of Study	Multifield	704,000	6	T	P	N
Modern Science Ontology	Multifield	369	6	O	P	N
Nature Subjects	Multifield	2,852	8	O	P	N
OpenAIRE's Field of Science Taxonomy	Multifield	50,000	6	T	P	N
OpenAlex Topics	Multifield	4,798	4	T	M	N
Research Fields, Courses and Disciplines (ASRC)	Multifield	1,061	3	T	M	N
Science Metrix Classification	Multifield	199	3	T	M	N
Socio-Economic Objective (ANZSRC)	Multifield	1,974	3	T	M	N
Subject Resource Application Ontology	Multifield	435	8	O	P	Y
UNESCO Thesaurus	Multifield	4,482	6	O	P	Y
Web of Science Categories	Multifield	254	1	L	–	N

National Library of Medicine classification, and the *Biomedical Ontologies from BioPortal*, although the last expands toward Biomedicine and hence more into the biological basis of health and disease. "Computer Science" can rely on three KOSs: the *Computer Science Ontology*, the *ACM Computing Classification Scheme*, and *Computer Science Subject Headings from Wikipedia*. In the field of "Biology," we identified two KOSs: the *Open Biological and Biomedical Ontology* and *EDAM*. However, *EDAM* focuses more on "Bioinformatics," describing the applied side of "Biology" with "Computer Science" tools. In the field of "Geology," we found the *GeoRef Thesaurus* and the *U.S. Geological Survey Library Classification System*. Economics is also covered by two KOSs: the *Journal of Economic Literature Classification System* and the *STW Thesaurus for Economics*. "Physics" can rely on the *Physics and Astronomy Classification Scheme* and *Physical Subject Headings*.

On the other hand, four KOSs were not directly assigned to one of the standard 19 broad fields, as they focus on more specialized areas. Specifically, we associated *Agrovoc* to "Agriculture," which is a subarea of the "Environmental Science"; *Biomedical Ontologies from BioPortal* to "Biomedicine," which covers both "Biology" and "Medicine"; *EDAM* to "Bioinformatics", which includes both "Biology" and "Computer Science"; and *TheSoz* to "Social Science," which is typically considered a superarea of "Economics," "Sociology," "Psychology," and "Political Science."

We were unable to identify a single-field KOS for the following seven fields: “History,” “Political Science,” “Environmental Science,” “Material Science,” “Geography,” “Sociology,” and “Business.” We reached out to a number of professors and domain experts in such fields, who confirmed this finding and mentioned that they usually rely on generic multifield KOSs, such as the *Dewey Decimal Classification* and the *Library of Congress Classification*. To the best of their knowledge, KOSs for such fields are yet to be developed.

In the category of multifield KOSs, we observed a substantial diversity in topic coverage. Figure 3 outlines the various fields covered by each KOS. For each system, the green fields represent the primary areas of focus, the blue fields indicate secondary areas with only a limited number of research topics, and the orange fields are those that are merely mentioned without any specific subareas. Notably, only five of these KOSs, emphasized in bold within the table, consistently cover all the 19 top-level research areas presented in Section 3.

4.2. Structure

This section analyzes the KOSs based on the previously defined structural features, focusing on their size, depth, type, hierarchical organization, and the presence of related terms.

4.2.1. Type of KOS

Table 1 (column “Sy”) indicates the category of each system using the following designations: **T** (taxonomy), **O** (ontology), **U** (thesaurus), and **L** (term list). Taxonomies (23 KOSs) and ontologies (18 KOSs) demonstrated the highest prevalence. In contrast, term lists and thesauri were less represented, with only one (*Web of Science Categories*) and three KOSs (*UMLS, IEEE Thesaurus*, and *GeoRef Thesaurus*), respectively.

This scenario suggests a clear division between the two approaches to categorizing research topics. On one side, some communities prefer to use straightforward hierarchical taxonomies, often encoded in very simple formats or structured documents. On the other side,

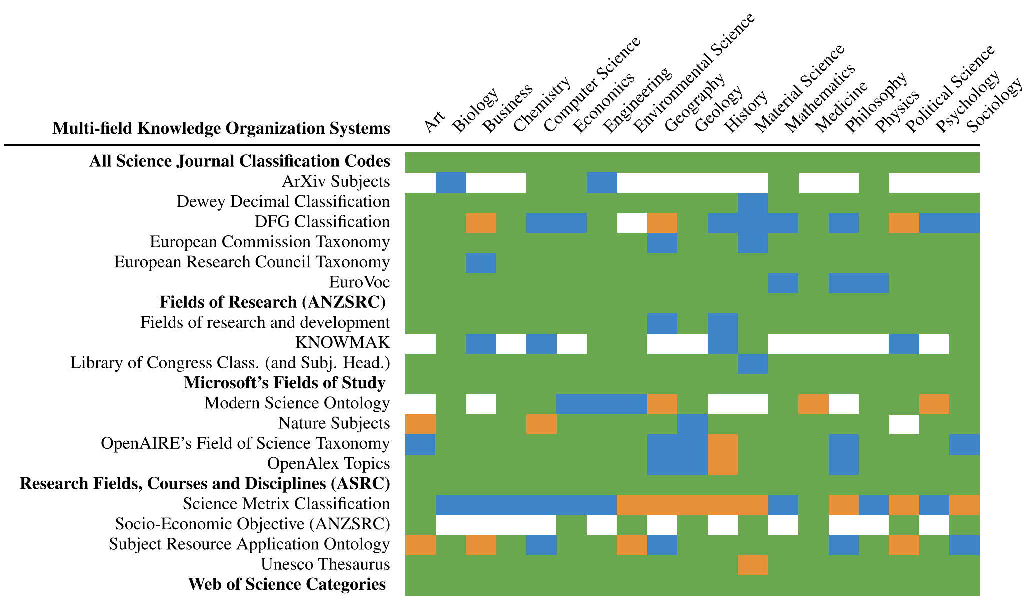


Figure 3. Coverage of the multifield KOSs. In green the main fields, in blue the minor ones, only partially represented, and in orange the fields that are just mentioned. In **bold** are the KOSs that consistently cover all research fields.

other communities embrace the richer expressivity of ontologies, which enable detailed descriptions of research topics, the relationships between them (e.g., causal, contributory, part/whole, and ancestral), and constraints (Kendall & McGuinness, 2019). Thesauri may be less common because they are more complex than hierarchical taxonomies but lack the full expressive power of ontologies.

4.2.2. Number of concepts

The number of concepts within KOSs varies widely, spanning from smaller systems like *Fields of Research and Development* (48 concepts) to vast ontologies like the *Biomedical Ontologies from BioPortal* (13 million concepts). Fourteen KOSs contain fewer than 1,000 concepts, 17 have between 1,000 and 10,000 concepts, nine include 10,000 to 100,000 concepts, and five have more than 100,000 concepts.

The median number of concepts within the analyzed KOSs is approximately 4,700. Sixteen single-field KOSs exceed this median in concept count, while 16 of the multifield KOSs contain fewer concepts. This pattern suggests a trend: Single-field KOSs tend to be larger and more specialized, likely to capture the intricacies within their specific domains. In contrast, multifield KOSs seem to be designed to offer a broader overview across various fields, often resulting in a smaller number of concepts.

The six multifield KOSs that exceed the median number of concepts, thereby providing a more granular representation of topics, are *OpenAlex Taxonomy* (4,798 concepts), *EuroVoc* (7,423), *OpenAIRE's Field of Science Taxonomy* (50,000), *Dewey Decimal Classification* (60,000), *Library of Congress Classification* (467,000), and *Microsoft's Fields of Study* (704,000).

4.2.3. Depth

Similarly to the number of concepts, the depth of KOSs also displays considerable variety, ranging from 1 (*Web of Science Categories*) to 39 (*Open Biological and Biomedical Ontology*). Specifically, 15 KOSs have up to three levels, six feature between four and five levels, nine employ exactly six levels, three use between seven and 10 levels, and 11 extend beyond 10 levels, as illustrated in Figure 4.

The median depth stands at six, with 18 of the 22 multifield KOSs falling below this threshold. This finding further supports the notion that the majority of multifield systems aim to provide a broad overview across diverse research fields.

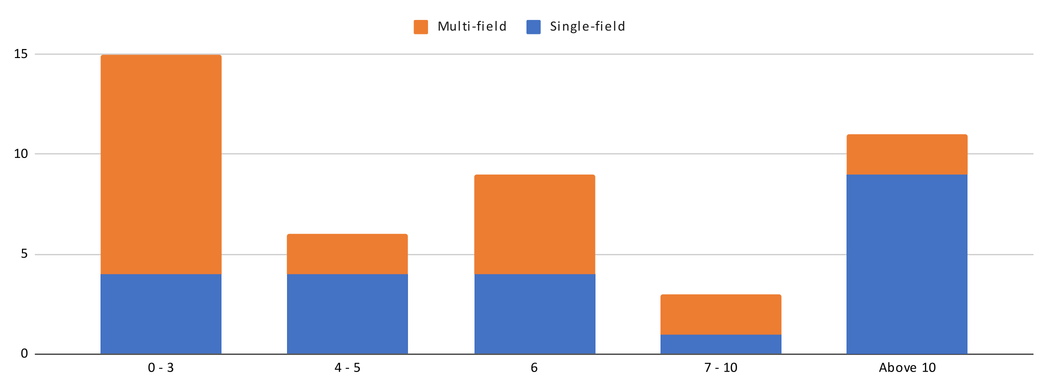


Figure 4. Depth distribution for single-field and multifield KOSs.

Table 2. Distribution of KOSs according to the type of hierarchy and their coverage (single or multifield). The * value does not account for *Web of Science Categories*, which is nonhierarchical

	Monohierarchical	Poly-hierarchical	Total
Single-field KOSs	9	14	23
Multifield KOSs	11	10	21*
Total	20	24	

4.2.4. Kind of hierarchy

Column "Hr" in Table 1 reveals a near-equal distribution between poly-hierarchical (24) and monohierarchical (20) KOSs. Only one KOS (*Web of Science Categories*) is considered non hierarchical as it is a flat list of terms. An analysis of the 23 single-field KOSs reveals that poly-hierarchical structures are more prevalent than monohierarchical ones (see Table 2). In contrast, multifield systems display a nearly balanced distribution, with 11 monohierarchical and 10 poly-hierarchical structures. This trend suggests that single-field KOSs tend to favor poly-hierarchical structures, potentially due to the need to represent finer-grained research topics that often stem from multiple parent topics.

4.2.5. Related terms

Table 1 (column "RT") reveals that the 21 KOSs incorporating related terms are mainly ontologies and thesauri. This aligns with the inherent capacity of these structures to express associative relationships (e.g., related terms). Interestingly, 15 of them are single-field KOSs.

Noteworthy exceptions are the *Mathematics Subject Classification* and the *Dewey Decimal Classification*. Although the *Mathematics Subject Classification* is traditionally considered a taxonomy (Dunne & Hulek, 2020), it also includes related terms. This functionality is enabled on the zbMATH website[18], where the KOS is displayed with a "see also" anchor, without adherence to the ANSI/NISO standard (ANSI/NISO Z39.19-2005(R2010); National Information Standards Organization, 2010). Similarly, the *Dewey Decimal Classification* uses the "see-also" references for synonyms and related terms (Mitchell, 2001).

4.3. Curation

This section explores the creation of KOSs, focusing on their generation methodologies, distribution formats, licensing, update frequency, available languages, and curators.

4.3.1. Generation

As previously discussed, the generation of KOSs can be categorized as manual, automated, semiautomated, or community based.

The majority (32) of the analyzed KOSs are developed manually. Developing these large knowledge bases manually is typically both time-consuming and very expensive.

Six KOSs have already adopted semiautomatic methodologies: *Biomedical Ontologies from BioPortal, KNOWMAK, Microsoft's Fields of Study, OpenAIRE's Field of Science Taxonomy, Subject Resource Application Ontology*, and *OpenAlex Topics*. For instance, *OpenAlex Topics* combines manually curated concepts from the *All Science Journal Classification Codes*,

[18] zbMATH: https://zbmath.org/classification.

with over 4,500 new research topics automatically identified through citation clustering. The approach first involved clustering papers based on citation patterns to form thematically related groups. These groups were then labeled using large language models and subsequently integrated with the existing concepts in the *All Science Journal Classification Codes* (OpenAlex, 2024).

Only two KOSs are generated using a fully automated pipeline, both within the field of computer science: the *Computer Science Ontology* and *Computer Science Subject Headings from Wikipedia*. The *Computer Science Ontology* was generated using Klink-2 (Osborne & Motta, 2015), which processed 16 million scientific publications. Klink-2 identifies relationships between topics by analyzing various indicators, including co-occurrence patterns, temporal distributions, and label similarity. The *Computer Science Subject Headings from Wikipedia* was created through an automated approach aimed at enhancing and refining the "Computer Science" branch of the Wikipedia category system. This approach integrates community detection, machine learning, and manually developed heuristics to identify and incorporate additional topics from Wikipedia articles (Han, Yang et al., 2020).

Finally, five KOSs leverage direct community expertise for their construction: *Open Biological and Biomedical Ontology, Medical Subject Headings, Unified Medical Language System, PhilPapers Taxonomy*, and *Physical Subject Headings*. They use various technologies to facilitate collaboration and enable researchers to suggest modifications to the KOSs. For instance, *Physical Subject Headings* encourages researchers to propose changes through GitHub issues, while *Medical Subject Headings* requests researchers to submit cases on their portal to suggest new terms, alterations, or corrections to the tree structure. *PhilPapers Taxonomy* consults experts in relevant areas, gathers insights from forum discussions, and incorporates feedback provided to the editors.

4.3.2. Distribution formats

Table 3 details the diverse formats in which KOSs are released. These range from PDF and HTML to more machine-readable options like CSV (and Excel), TSV, MARC, and RDF[19] (Resource Description Framework), an open standard established by the World Wide Web Consortium[20] (W3C). Several KOSs are released in multiple formats to cover different use cases.

The majority of KOSs (27) are accessible in HTML format, primarily through their providers' websites. In this format, the KOS is essentially a webpage listing and connecting concepts. While this is human readable, the underlying data lack the structured organization necessary for direct and automated processing by computer systems. Several KOSs are available in machine-interpretable formats: 17 utilize the RDF standard, and 10 are available in CSV format. Finally, six KOSs are exclusively available in PDF format, which complicates their integration with other KOSs and hinders automatic analysis.

4.3.3. License

The majority of the KOSs under analysis (26) employ open licenses from Creative Commons (CC), Open Data Commons (ODC), Open Database License (ODbL), or MIT licenses. Their openness varies considerably, as systems like *Physical Subject Headings* (CC0 1.0) are highly open, while others, such as *TheSoz* (CC BY-NC-ND 3.0), impose slightly stricter terms. Nonetheless, the CC BY license remains the most prevalent.

[19] Resource Description Framework: https://www.w3.org/RDF.

[20] World Wide Web Consortium: https://www.w3.org.

Table 3. Analysis of the formats in which the different KOSs are currently released. Single-field KOSs are in the upper part and multifield KOSs are in the lower part

Knowledge organization system	RDF	CSV/Excel	HTML	PDF	Other formats
Agrovoc Thesaurus	x		x		
Art and Architecture Thesaurus	x		x		
EDAM	x	x	x		TSV
Open Biological and Biomedical Ontology	x		x		OBO
Biomedical Ontologies from BioPortal	x	x			OBO
ChemOnt			x		OBO, JSON
ACM Computing Classification Scheme	x		x	x	
Computer Science Ontology	x		x		
Computer Science Subject Headings from Wikipedia		x			
Journal of Economic Literature			x		XML
STW Thesaurus for Economics	x				
IEEE Thesaurus				x	
GeoRef Thesaurus				x	
U.S. Geological Survey Library Classification System				x	
Mathematics Subject Classification			x	x	
Medical Subject Headings	x		x		MARC
National Library of Medicine classification			x	x	
Unified Medical Language System			x		RRF, ORF, SQL
PhilPapers Taxonomy			x		
Physical Subject Headings	x		x		
Physics and Astronomy Classif. Scheme			x		
PsycInfo and PsycTests Classification Systems			x	x	
TheSoz	x				
All Science Journal Classification Codes			x		
ArXiv Subjects			x		
Dewey Decimal Classification			x	x	
DFG Classification			x	x	
European Commission Taxonomy				x	
European Research Council Taxonomy				x	
EuroVoc	x	x	x		MARC

Table 3. (*continued*)

Knowledge organization system	RDF	CSV/Excel	HTML	PDF	Other formats
Fields of Research (ANZSRC)		x	x		
Fields of research and development				x	
KNOWMAK	x				JSON
Library of Congress Class. (and Subj. Head.)	x				MARC
Microsoft's Fields of Study					TSV
Modern Science Ontology	x				
Nature Subjects		x			
OpenAIRE's Field of Science Taxonomy			x		JSON
OpenAlex Topics		x			JSON
Research Fields, Courses and Disciplines (ASRC)		x	x		
Science Metrix Classification		x			
Socio-Economic Objective (ANZSRC)		x	x		
Subject Resource Application Ontology	x				
UNESCO Thesaurus	x		x		
Web of Science Categories			x		

Eleven KOSs are freely accessible online but maintain copyright restrictions, meaning that they can be browsed but cannot be downloaded, modified, or redistributed without explicit permission from the copyright holder.

Two KOSs are free from copyright: *Medical Subject Headings* and *National Library of Medicine classification*. Four KOSs are copyrighted and inaccessible online. For instance, both the *Dewey Decimal Classification* and the *Library of Congress Classification* necessitate licensing fees even for browsing purposes.

Finally, *Open Biological and Biomedical Ontology* and *Biomedical Ontologies from BioPortal* represent broader initiatives that integrate multiple ontologies. The licensing terms for these efforts are contingent on the individual ontologies they incorporate. For the former, its mission dictates that all incorporated ontologies must be openly accessible under licenses such as CC BY 3.0, CC BY 4.0, or CC0 1.0, ensuring unrestricted use. In contrast, BioPortal presents a more diverse licensing landscape. Indeed, while many ontologies are openly available, some have specific terms of use set by their providers.

4.3.4. Frequency of updates

Determining the frequency of updates proved to be a significant challenge, as providers rarely disclose their updating schedules explicitly. In some instances, we were able to infer updating patterns by analyzing the dates of previous releases.

Eight KOSs are continuously updated and maintained, with new revisions produced monthly or more frequently. These include the *Library of Congress Subject Headings*, the *Art & Architecture Thesaurus*, the *UNESCO Thesaurus, TheSoz*, the *Dewey Decimal Classificaion*, the *Open Biological and Biomedical Ontology*, the *PhilPapers Taxonomy*, and *Agrovoc*.

Three systems are updated regularly, although less frequently, with revisions taking place twice a year: *Unified Medical Language System, National Library of Medicine classification*, and *EuroVoc*.

Seven KOSs are updated once a year: the *STW Thesaurus for Economics*, the *Medical Subject Headings*, the *Computer Science Ontology, OpenAlex Topics*, the *IEEE Thesaurus, OpenAIRE's Field of Science Taxonomy*, and the *Subject Resource Application Ontology*. Five KOSs receive less frequent updates (roughly every 10–15 years); these include the *ACM Computing Classification System*, the *Mathematics Subject Classification*, the *Science Metrix Classification, Socio-Economic Objective*, and the *Fields of Research*.

Finally, three KOSs are no longer actively maintained but continue to be used by their respective communities: *Research Fields, Courses and Disciplines, Microsoft's Fields of Study*, and the *Physics and Astronomy Classification Scheme*.

There are many factors that can influence the frequency of updates of a KOS. One factor is related to the discipline and its evolution pace. For instance, the field of "Medicine" is a fast-advancing field, and for this reason, there is a new version of *Medical Subject Headings* released every year.

Another factor influencing the frequency of updates is the depth (i.e., specificity) of a KOS. For instance, *Agrovoc Thesaurus*, whose depth equals 14, requires more frequent updates to include new emerging topics and readjust the hierarchical structure due to epistemological changes. For this reason, *Agrovoc Thesaurus* receives monthly updates. In contrast, the *Fields of Research (ANZSRC)* consists of only three levels, and because the modeled concepts can be considered quite generic, it is reasonable to assume that their structure will uphold over a relatively longer timespan. Indeed, the *Fields of Research (ANZSRC)* received its most recent update in 2020, after 12 years.

We also investigated whether the frequency of updates is related to the type of system, as shown in Figure 5. Our analysis revealed that among the 18 KOSs updated within a year, the majority were ontologies (11), followed by taxonomies (five), and thesauri (two). In contrast, the eight KOSs that were either updated every 10 years or discontinued were taxonomies. This suggests that while taxonomies might be effective initially for organizing subjects hierarchically, they may become increasingly challenging and expensive to maintain over time due to the growing complexity and unique characteristics of the subjects they represent. On the other hand, ontologies, often implemented in flexible formats such as OWL and RDF, may be easier to maintain.

4.3.5. Last update

Twenty systems received their most recent update in 2023 or later, while 13 were last updated between 2020 and 2022. This indicates that within the last five years, 33 KOSs have been revised to reflect the evolving nature of their respective fields. Of these, 17 are single-field systems and 15 are multifield. The remaining 12 systems either received their last update before 2020 or their latest release date could not be determined.

Notably, key disciplines such as Chemistry (with *ChemOnt*, last updated in 2016) and "Geology" (with the *U.S. Geological Survey Library Classification System*, last updated in

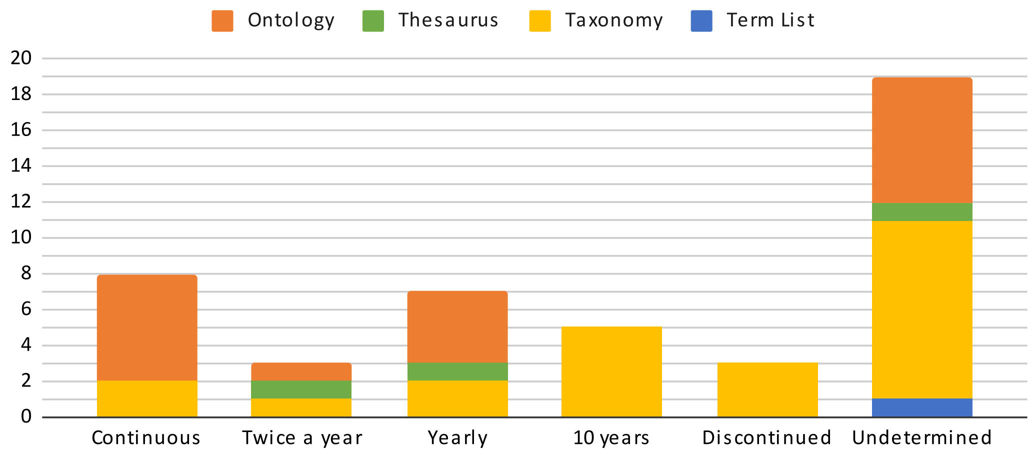


Figure 5. Frequency of update of KOSs in relation to their type.

2000, and the *GeoRef Thesaurus*, last updated in 2008) can only rely on possibly outdated KOSs. This is potentially a significant issue, as the outdated representations in these disciplines can hinder the dissemination of modern research efforts that address topics not fully covered by these KOSs.

4.3.6. Languages

Our exclusion criteria ensured that all KOSs included in this study had at least one version available in English. Specifically, 35 KOSs are exclusively available in English. In contrast, 10 KOSs provide research topics in multiple languages, reflecting their intended applications and jurisdictional requirements. For instance, *EuroVoc*, developed by the Publications Office of the European Union, supports 23 languages, enabling interoperability across European digital libraries. Table 4 lists the 10 multilingual KOSs, detailing the number of languages they

Table 4. This table presents a comparison of 10 multilingual KOSs, detailing the number of languages they cover, whether language coverage is uniform, their scope coverage (Single-field or Multifield), type of system (**Sy** = **O**ntology, **T**axonomy, and Thesa**U**rus), number of concepts (#Concepts), and depth

Knowledge organization system	Languages	Uniform	Scope Coverage	Sy	#Concepts	Depth
Art and Architecture Thesaurus	167		Single-field	O	58,000	13
Agrovoc Thesaurus	42		Single-field	O	41,000	14
Dewey Decimal Classification	30		Multifield	T	60,000	14
Unified Medical Language System	28		Single-field	U	3.3 million	30
Science Metrix Classification	26	Yes	Multifield	T	199	3
EuroVoc	23	Yes	Multifield	O	7,439	6
Fields of research and development	6	Yes	Multifield	T	48	2
UNESCO Thesaurus	5	Yes	Multifield	O	4,482	6
TheSoz	4		Single-field	O	8,223	6
STW Thesaurus for Economics	2	Yes	Single-field	O	14,000	13

support and other relevant characteristics to provide a comprehensive overview of their features.

The 10 multilingual KOSs exhibit substantial variability in the number of supported languages. Some provide broad language support, such as the *Science Metrix Classification* (26 languages) and the *Agrovoc Thesaurus* (42 languages). Notably, the *Art and Architecture Thesaurus* demonstrates exceptional linguistic inclusivity with its impressive coverage of 167 languages. In contrast, other KOSs support only a few languages, such as the *STW Thesaurus for Economics* (German and English), the *UNESCO Thesaurus* (five languages), and the *Fields of research and development* (six languages).

Furthermore, KOSs often vary in the number of concepts available in different languages. Typically, non-English versions only provide a partial representation of the domain described by the English version. For instance, as shown in Figure 6, *Agrovoc* provides good coverage in English, French, Turkish, Spanish, Arabic, and a few other languages. However, the number of available concepts is significantly limited in languages such as Estonian, Burmese, Khmer, and Greek. Naturally, achieving uniform coverage poses greater challenges for KOSs with both a high volume of concepts and extensive multilingual support. Only a few smaller KOSs, such as *Eurovoc, UNESCO Thesaurus, Fields of Research and Development, Science Metrix Classification*, and *STW Thesaurus for Economics*, have managed to maintain consistent representation across all supported languages, as reported by the column "Uniform" in Table 4. Notably, *TheSoz* provides near-uniform coverage in English, German, and French, while its Russian counterpart remains under development.

Finally, an analysis of the 10 multilingual KOSs reveals that the majority (six) are ontologies, while three are taxonomies, and one is a thesaurus. This trend likely reflects the inherent capacity of ontologies to effectively manage concept labels across multiple languages (Montiel-Ponsoda, De Cea et al., 2011).

4.3.7. Providers and maintainers

We categorized the providers of all knowledge organization systems into five distinct groups: Publishers & Information Service Providers, Funding Bodies & Government Agencies & Research Councils, & Policy Makers, Research Institutes & Universities, National Libraries, and Open Initiatives & Consortia.

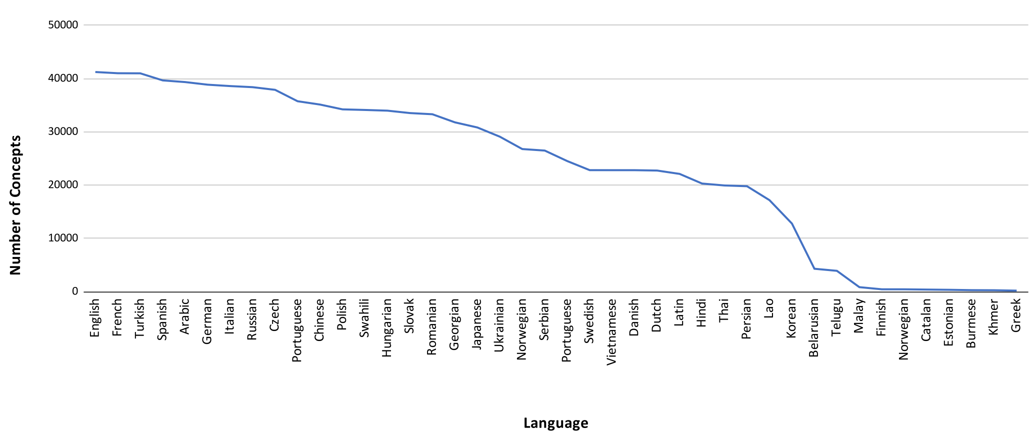


Figure 6. Distribution of terms per language in the *Agrovoc Thesaurus*. Produced with term counts available on https://agrovoc.fao.org/browse/agrovoc (Agrovoc v.2024-04, April 2024).

Publishers & Information Service Providers represent the most prevalent category, with 16 KOSs, including notable members such as Elsevier, Springer Nature, the Institute of Electrical and Electronics Engineers (IEEE), the Association for Computing Machinery (ACM), Clarivate, and the American Psychological Association (APA). Indeed, KOSs are essential for publishers as they need to organize, manage, and deliver content effectively in order to enhance discoverability, accessibility, and user engagement (Salatino et al., 2019).

The second-largest group comprises Funding Bodies, Government Agencies, Research Councils, & Policy Makers. This category includes 12 organizations such as the European Commission, UNESCO, the Food and Agriculture Organization (FAO), the Australian Research Council, the New Zealand Ministry of Business, and the Organisation for Economic Co-operation and Development (OECD). In this context, KOSs are crucial for streamlining operations, promoting transparency, and enhancing decision-making.

Research Institutes and Universities, including The Open University (United Kingdom), University of Sheffield (United Kingdom), University of Illinois at Urbana-Champaign (United States), University of Alberta (Canada), Athena Research Center (Greece), and The Getty Research Institute (United States) produced eight KOSs.

National Libraries, including the US National Library of Medicine, the Library of Congress, the German National Library of Economics (ZBW), and the German National Library of Science and Technology (TIB) provided seven KOSs.

Finally, only two KOSs are supported by Open Initiatives and Consortia: the Open Biological and Biomedical Ontology (OBO) Foundry (*Open Biological and Biomedical Ontology*) and FAIRsharing (*Subject Resource Application Ontology*).

Our analysis showed that Funding Bodies & Government Agencies & Research Councils, & Policy Makers primarily developed multifield KOSs (nine out of 12), likely due to their interdisciplinary focus. In contrast, Research Institutes & Universities predominantly created single-field KOSs (six out of eight), possibly because research teams tend to specialize in a specific area. The remaining organizational categories demonstrated a balanced approach in developing both single and multifield KOSs.

4.4. External Links

Nineteen KOSs provide links to external knowledge bases. This is typically done by defining mappings that indicate that two entities in different representations refer to the same real-world object or concept. They typically rely on well-known semantic relationships such as `owl:same_as`, `skos:exactMatch`, and `skos:closeMatch`. As an example, the concept "sunflowers" available in *Agrovoc Thesarus*[21] has a `skos:closeMatch` with sunflower in *Eurovoc*[22].

The KOSs under analysis typically connect either to general knowledge graphs (Peng, Xia et al., 2023), such as Wikidata[23] or DBpedia[24], or to other KOSs. Such interconnections are generally beneficial, as they create a network of related resources, providing diverse perspectives on research topics and facilitating the creation of novel resources.

[21] Sunflowers (plural) in *Agrovoc Thesarus:* https://aims.fao.org/aos/agrovoc/c_aad037e4.

[22] Sunflower (singular) in *Eurovoc:* https://eurovoc.europa.eu/4472.

[23] Wikidata: https://www.wikidata.org.

[24] DBpedia: https://www.dbpedia.org/.

Wikidata stands out as the most externally linked knowledge graph as it is reached by concepts from *Agrovoc Thesaurus, Computer Science Ontology, STW Thesaurus for Economics, EuroVoc, Open Biological and Biomedical Ontology*, and the *Library of Congress Class. (and Subj. Head.)*. Conversely, *Agrovoc Thesaurus, ChemOnt, STW Thesaurus for Economics*, and *EuroVoc* are the KOSs with the highest number of externally connected knowledge bases. For instance, the *Agrovoc Thesaurus* is highly interconnected, linking its concepts to *EuroVoc*, Wikidata, DBpedia, the *UNESCO Thesaurus*, and the *Library of Congress Subject Headings (LCSH)*. Similarly, *EuroVoc* is linked not only with the *Agrovoc Thesaurus*, but also with *TheSoz, LCSH, UNESCO, STW, MeSH*, Wikidata, and several other knowledge organization systems.

Thirteen out of the initial 19 earn a five-star rating within the Linked Open Data deployment scheme proposed by Sir Tim Berners-Lee in 2010[25]. Their extensive interconnections, facilitated by RDF standards such as OWL and SKOS, establish a rich network of information with related KOSs. Adhering to the five-star scheme[26] optimizes discoverability, reusability, and the potential for integration into the broader knowledge web, fostering innovative applications and the development of more comprehensive knowledge organization systems.

4.5. Usage

A KOS can be widely adopted across various applications, such as organizing digital libraries, enhancing metadata, and ensuring interoperability between different data systems. For example, the *Fields of Research (ANZSRC)* is utilized by Dimensions.ai to organize research metadata, by Figshare to manage research repositories, and by the Australian and New Zealand governments to measure and analyze research and experimental development.

Our analysis showed that KOSs are being employed to organize four main categories of resources: *digital libraries*, which contain research articles, policy documents, grant proposals, patents, and other kinds of digital documents; *research repositories*, which contain research data, code, research protocols, models, and any other kinds of research artefacts; *bibliographic databases*, which organize metadata about publications and research artefacts; and *physical libraries*, which contain printed books and periodicals, as well as other media.

Twenty-nine KOSs are currently being employed to organize digital libraries including ACM Digital Library, IEEE Digital Library, MEDLINE/PubMed, Scopus, Nature, EconLit, Physical Review, and Mathematical Reviews. Ten KOSs are employed to organize research outputs like *Modern Science Ontology* for the Open Research Knowledge Graph, *ChemOnt* for DrugBank, T3DB, ChEBI, LIPID MAPS, *TheSoz* for the Social Science Research Project Information System in Germany, *EDAM* for bio.tools and Training eSupport System (TeSS), *OpenAIRE's Field of Science Taxonomy* for OpenAIRE, and *Subject Resource Application Ontology* for FAIRsharing. Ten KOSs are employed to organize bibliographic databases, including Web of Science, Dimension.ai, Microsoft Academic Graph, APA PsycInfo database, the International System for Agricultural Science and Technology, and GeoRef database. Finally, the *National Library of Medicine classification, Dewey Decimal Classification*, and *Library of Congress Classification* are mainly being employed for physical libraries.

Various KOSs are also utilized across a wide array of initiatives to directly advance research and create additional knowledge and tools. For instance, the *Computer Science Ontology* has

[25] Linked Data: https://www.w3.org/DesignIssues/LinkedData.html.

[26] Linked Open Data 5 Star: https://www.w3.org/DesignIssues/LinkedData.html.

been employed for exploring and analyzing scholarly data through platforms like Rexplore (Osborne, Motta, & Mulholland, 2013), ScholarLensViz (Löffler, Wesp et al., 2020), ConceptScope (Zhang, Chandrasegaran, & Ma, 2021), and VeTo (Vergoulis, Chatzopoulos et al., 2020). It has also been instrumental in generating knowledge graphs, such as AIDA KG (Angioni, Salatino et al., 2022b) and CS KG (Dessì, Osborne et al., 2022), as well as in recommending video lessons (Borges & dos Reis, 2019). However, thoroughly analyzing these specific applications would require a dedicated survey, as exemplified by Jing's usage analysis of the *Unified Medical Language System* (Jing, 2021).

5. CHALLENGES AND FUTURE DIRECTIONS

In this section, we present the main challenges and future directions of this field. In Section 5.1, we focus on what we consider the most important challenge in this field: the creation of a comprehensive and granular KOS representing all academic disciplines. Section 5.2 discusses the most important limitations and challenges regarding the integration of KOSs. We then present several other important challenges characterizing this space, including: expanding the coverage of different languages (Section 5.3); reconciling expert disagreements (Section 5.4); assessing the quality of KOSs (Section 5.5); handling ambiguous labels (Section 5.6); improving the generation and frequency of updates (Section 5.7); and developing scalable and accurate approaches for item classification (Section 5.8).

5.1. Towards a Comprehensive and Granular Representation of All Academic Disciplines

An important limitation of the KOSs described in this work lies in their high fragmentation. On one side, we have a good number of multifield systems that cover multiple academic areas, but tend to be quite shallow and miss many research topics. On the other, we have a plethora of single-field KOSs, typically offering a much more granular representation of scholarly knowledge. However, there is no KOS that is simultaneously **comprehensive** (i.e., it encompasses all the main academic fields); **granular** (i.e., it represents all the specific research areas that are typically used by researchers to refer to their work); **maintained** (i.e., it is constantly updated to reflect the latest developments); and **open** (i.e., it can be used freely, without restrictions).

We argue that such a resource would be transformative in this field and allow organizing content across various digital libraries, thereby enhancing interoperability; facilitating the retrieval and analysis of research outputs (e.g., articles, patents, project reports) and agents (e.g., researchers, organizations); enabling the monitoring of research, development, and innovation activities; ensuring data quality and compatibility across research institutions; and supporting evidence-informed policymaking.

A potential approach to creating this resource is to adopt one or more multifield KOSs to represent broad research areas and then integrate several single-field KOSs to effectively cover specific disciplines. As a first step, this section will explore which KOSs could be integrated to develop a comprehensive system. In Section 5.2, we will discuss current methods for interlinking multiple KOSs.

5.1.1. Multifield KOSs that may support the creation of a general KOS

It is useful to analyze to what extent the current multifield KOSs fulfill the requirements defined in the previous session. Figure 7 presents an evaluation of the 22 multifield KOSs examined in this study, considering four key features: comprehensiveness, depth, frequency of updates, and openness. Each of these categories is characterized using a color-coded

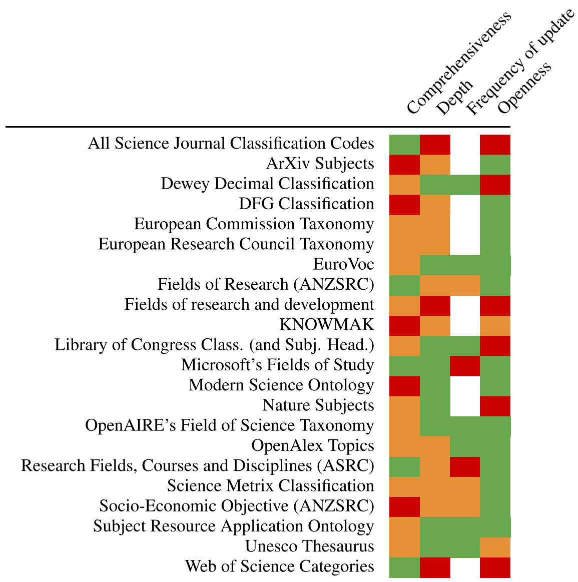


Figure 7. Analysis of KOSs according to their comprehensiveness, depth, frequency of update, and openness. Features of each KOS are evaluated using a color-coded system: ■ green indicates a strong performance, ■ orange signifies an acceptable performance, and ■ red reflects poor performance. A feature is blank when the information is not available.

system to indicate performance levels: strong (green), acceptable (orange), or poor (red). The criteria for assigning these performance levels are as follows.

For the comprehensiveness aspect, colors are assigned based on how well the KOSs cover the various fields, as detailed by Figure 3. Specifically, green indicates comprehensive coverage, orange denotes partial coverage, and red signifies that the KOSs cover only a limited set of fields. In the depth column, the color coding is as follows: Red indicates a depth of 2 or less (low or nonexistent), orange represents a depth between 3 and 5 (medium), and green signifies a depth of 6 or higher (high). In the frequency of update column, red indicates that the KOS is discontinued with no future updates planned, orange is used for slow updates (i.e., 5–10 years), and green denotes more frequent updates (i.e., within 5 years). If such information is unavailable, it is marked in white. Last, in the openness column, red indicates that the KOS is copyrighted and requires a special license for use, orange denotes that the KOS is available online and can be used with certain restrictions, and green signifies a fully open resource with no usage limitations.

Only five KOSs include all 19 broad fields of study: *All Science Journal Classification Codes, Fields of Research (ANZSRC), Research Fields, Courses and Disciplines (ASRC), Microsoft's Fields of Study*, and *Web of Science Categories*. However, none of these KOSs fulfills all the other requirements. *All Science Journal Classification Codes* and *Web of Science Categories* are very shallow and not open. The *Fields of Research, Research Fields, Courses and Disciplines (ASRC)*, and *Microsoft's Fields of Study* offer more depth compared to the previous two, but they were discontinued. *Fields of Research (ANZSRC)* is limited in depth, with many specific categories still overly broad, and updates are infrequent, often taking several years to implement. The *Dewey Decimal Classification* and *Library of Congress Subject Headings* offer extensive coverage across nearly all areas, along with substantial depth and frequent updates. However, their openness remains an issue, as users need to pay subscription fees to use them.

In conclusion, the current state-of-the-art multifield and open KOSs do not yet provide all the essential features necessary to serve as a robust foundation for a comprehensive general KOS. However, there are a few promising candidates that, if integrated effectively, could provide a solid starting point. In particular, one potential research path could begin with *Eurovoc* as a multifield KOS, incorporating missing concepts from specialized single-field KOSs like *Mathematics Subject Classification, PhilPapers Taxonomy*, and *Physical Subject Headings*.

5.1.2. Academic fields that are not currently covered by specific KOSs

Several disciplines lack dedicated KOSs and are only superficially addressed by multifield systems, making it challenging to categorize documents with the necessary precision. In particular, we were unable to identify any KOS that offers a good characterization of seven research fields: "History," "Political Science," "Environmental Science," "Material Science," "Geography," "Sociology," and "Business." In the fields of "Political Science" and "Sociology," *TheSoz* could potentially be utilized. However, *TheSoz* mostly focuses on "Social Science" and does not fully cover these two disciplines. The absence of fine-grained representation for these seven major areas is notable and underscores the need for further development in this space. Automated methods for generating KOSs can offer a valuable solution in this regard, a topic we will further discuss in Section 5.7.

5.2. Integration of Multiple KOSs

In the following, we discuss future directions regarding the integration of multiple KOSs, which may lead to creating a more comprehensive and granular representation of research fields.

A few of the KOSs that we described in this survey are already interlinked. For instance, some concepts in the *Subject Resource Application Ontology* are mapped to concepts within *AgroVoc* and *EDAM*, as well as to a number of ontologies available in *OBO*. *AgroVoc*, on its turn, has some concepts mapped to the *UNESCO Thesaurus*, and the *Library of Congress Subject Headings*. However, several KOSs, like the *Mathematical Subject Classification* and the *Physical Subject Headings*, are not (yet) connected to any other KOS. Moreover, the existing mappings between KOSs are often incomplete. The integration of KOSs presents several challenges, which can be categorized into the following subchallenges:

- **Subchallenge 1:** Generating mappings between KOSs;
- **Subchallenge 2:** Adopting standard formats;
- **Subchallenge 3:** Developing tools for facilitating the integration of KOSs.

5.2.1. Generating mappings between KOSs

Generating links between KOSs is a complex task. In the academic domain, this is usually done by identifying that two subjects from different systems refer to the same concept and linking them with a relation, such as `owl:sameAs` and `skos:exactMatch`. This process can be either manual, automatic, or semiautomatic (Kalfoglou & Schorlemmer, 2003).

Manual approaches are convenient for small KOSs, but they typically require a lot of effort and high-level expertise and may suffer from scalability issues. In addition, manual integration can lead to the introduction of inconsistencies, especially for large KOSs (Erdogan, Erdem, & Bodenreider, 2010; Halper, Morrey et al., 2011; Solimando, Jiménez-Ruiz, & Guerrini, 2014).

Automatic approaches typically use a combination of similarity metrics, natural language processing, and machine learning (Declerck, 2013; Salatino, Thanapalasingam et al., 2020; Zapilko, Schaible et al., 2013). For instance, Salatino, Thanapalasingam et al. (2020) associated research topics to the corresponding DBpedia entity with the DBpedia Spotlight API (Daiber, Jakob et al., 2013). They fed the tool with artificial sentences listing the labels of the topic and of its direct sub- and supertopics, and then it returned the related DBpedia entities alongside the similarity score. Zapilko et al., (2013) mapped *TheSoz* to *Agrovoc* and DBpedia, using string similarity between terms and retaining the matches that had Levenshtein distance lower than a threshold. In the domain of KOSs of academic fields, we still typically rely on simple approaches, which mainly use lexical heuristics, and may lead to three potentially unintended consequences (Shvaiko & Euzenat, 2008; Slater, Gkoutos, & Hoehndorf, 2020; Solimando et al., 2014). First, the mapping might introduce new internal relationships between the entities of one system, and therefore modify inadvertently the description of the domain. Instead, the mapping should just enable the interaction across KOSs. Second, the automatic integration can introduce logical inconsistencies. Finally, the mapping might connect entities belonging to different contexts, indicating a potential mapping error (Jiménez-Ruiz, Grau et al., 2011). Some recent approaches address these limitations by relying on description logic (Dhombres & Bodenreider, 2016) or deep learning (Yip, Nguyen, & Bodenreider, 2019). For instance, Dhombres and Bodenreider (2016) developed an approach for mapping *Human Phenotype Ontology* (HPO) and the *Standardized Nomenclature of Medicine Clinical Terms* (SNOMED CT) using both lexical and logical approaches. The latter approach consists of using a representation based on description logic to compare the concepts. This method can mitigate the limitations of the lexical approach; however, not all KOSs are developed through description logic. Yip et al. (2019) developed a deep learning model with a Siamese recurrent architecture to identify synonyms across *UMLS* concepts. The model provides good results; however, it still requires more research and fine-tuning because it presents several false positives (matches together nonsynonyms) and false negatives (fails to identify synonyms). The community still needs to further investigate and develop these new solutions in practical settings.

Finally, in *Semiautomatic approaches*, domain experts analyze, correct, and give feedback on candidate mappings produced by automatic approaches (Salvadores, Alexander et al., 2013; Yip et al., 2019). It is worth mentioning two big endeavors in this space: *Unified Medical Language System* and the *Biomedical Ontologies from BioPortal*. The *Unified Medical Language System* is maintained by the US National Library of Medicine and currently integrates more than 200 vocabularies across different languages in the field of Medicine (Bodenreider, 2004). The US National Library of Medicine is willing to include new additional vocabularies as long as they meet the criteria for inclusion[27]. These include whether the vocabulary brings new or unique content, is actively maintained, and is available in a machine-readable format. The *Biomedical Ontologies from BioPortal* integrates more than 1,100 ontologies in the field of "Biomedicine". Differently from *UMLS*, registered users can submit new ontologies to the BioPortal without constraints.

In conclusion, generating new interconnections between KOSs—either manually or through automatic approaches—is still an open challenge.

[27] Unified Medical Language System inclusion evaluation criteria: https://www.nlm.nih.gov/research/umls/knowledge_sources/metathesaurus/source_evaluation.html.

5.2.2. Adopting standard formats

The format of a KOS has a great impact on our ability to interconnect it with other knowledge bases. As shown in Table 3, some KOSs are published in RDF[28] (i.e., Resource Description Framework), which is a World Wide Web Consortium[29] (W3C) standard and used for representing highly interconnected data, such as the *UNESCO Thesaurus*, and the *Medical Subject Headings*. Some other KOSs are published in CSV, PDF, or browsable through web pages (HTML).

Based on the 5-star deployment scheme for open data[30], among the possible publishing formats, RDF is the one that fosters better integration. A system published according to this standard consists of a set of RDF statements. Each statement is a three-part structure (also known as triple) which is the smallest irreducible representation for binary relationships, and it is expressed in the form of <*subject, predicate, object*> (Berners-Lee, Hendler, & Lassila, 2001). For instance, <*Social Sciences, narrower, Politics*> indicates that Social Sciences is a broader area of Politics. In this way, two entities (both subject and object) are linked via a predicate or verb. In addition, every part of a triple is individually addressable through unique URIs, such as: <*https://www.nature.com/subjects/social-sciences*, *https://www.w3.org/2004/02/skos/core#narrower*, *https://www.nature.com/subjects/politics*>. Such representation allows AI systems to interconnect, identify, disambiguate, and integrate data effectively.

In contrast, KOSs published in CSV or HTML are more challenging to treat because the relevant knowledge is not described in a structured format, and relations may be implicit or interpretable only in a wider context. Generating a representation of these KOSs in a standard format such as RDF is not a trivial task. For example, when working with CSV files, it is essential to understand the structure and data types of the columns (e.g., string, integer, float, date). Tools such as RML (Dimou, Vander Sande et al., 2014) can be used to automate the conversion process from CSV to RDF. However, this process still requires careful consideration of the data's format and structure to ensure accurate transformation. Similarly, KOSs published in HTML often have arbitrary structures, making it necessary to develop custom parsers to convert them into RDF. When KOSs are available only as PDF files, parsing and extracting the information becomes even more challenging.

In conclusion, the RDF format is adopted by only half (11 out of 23) of the single-field KOSs across the different disciplines, as shown in Table 3. Several significant academic fields (e.g., "Psychology," "Chemistry," "Geology," "Engineering," "Philosophy") do not yet rely on standard machine-readable formats, hindering the reuse and integration of their KOSs. In the literature, we can find a few research efforts to RDFy[31] KOSs. Examples include the RDFification of the *Mathematics Subject Classification*[32] and *IEEE Thesaurus*[33]. However, these RDF versions are not maintained by the original curators of the KOSs, making them difficult to update and reuse.

[28] Resource Description Framework: https://www.w3.org/RDF.
[29] World Wide Web Consortium: https://www.w3.org.
[30] 5-star Open Data: https://www.w3.org/DesignIssues/LinkedData.html.
[31] *RDFication* refers to the process of converting data into Resource Description Framework (RDF) format.
[32] MSC2020_SKOS (TIBHannover): https://github.com/TIBHannover/MSC2020_SKOS.
[33] ieee-taxonomy-thesaurus-rdf (The Open University): https://github.com/angelosalatino/ieee-taxonomy-thesaurus-rdf.

5.2.3. Developing tools for facilitating the integration of KOSs

The first KOSs were of relatively small size, such as the *Library Classification for Environmental Science*[34], developed by Plate (1966). Typically, their creation would involve sketching ideas on paper or writing down topics on sticky notes, which would then be arranged on a table to create a hierarchical structure (Motta, 1999). Today, we have access to a wide range of tools that streamline this process, enabling the creation of more structured and complex KOSs. For instance, the German Centre for Higher Education Research and Science Studies[35] (DZHW) used Trello to build the *Research Core Dataset*[36] (RCD), a classification of interdisciplinary research fields (Stiller, Trkulja et al., 2021). Trello[37] is a web-based project management tool implementing the Kanban system, and it allows multiple users to collaborate on the same board. To build *RCD*, the experts created a card for each subject and then arranged them over the board to create the hierarchical structure. However, building complex KOSs using Trello presents some limitations, as it does not allow users to nest cards on more than two levels. In addition, as Trello is not intended for this task, it does not provide any tool to handle disagreement among the experts due to their different backgrounds.

More advanced tools for building KOSs are Protégé[38], Semantic MediaWiki[39], VocBench3[40], and PoolParty[41]. Protégé is an open-source software developed by Stanford University to support developers in creating reusable ontologies and building knowledge-based systems (Musen, 2015). Its graphical interface offers a range of functionalities for browsing and editing ontologies. While the original Protégé software lacks native collaborative features, its cloud-based counterpart, WebProtégé, enables multiple users to work simultaneously on ontology development (Tudorache, Nyulas et al., 2013).

Semantic MediaWiki is an open-source extension of MediaWiki, the engine that runs underneath Wikipedia (Vrandečić & Krötzsch, 2009). Semantic MediaWiki provides a stable, powerful, and scalable environment, enabling users to browse and collaboratively edit ontologies. It also provides facilities for rating pages and users.

VocBench3 is an open-source web application that allows users to create, manage, and share ontologies and thesauri (Stellato, Fiorelli et al., 2020). It provides a user-friendly interface for editing and managing RDF data, supports collaborative editing and version control, enabling tracking of changes and the ability to revert to prior versions.

PoolParty is a suite that supports the creation and maintenance of taxonomies, ontologies, knowledge graphs, and semantic search applications (Schandl & Blumauer, 2010). Similarly to WebProtégé, it allows users to browse and work collaboratively. Other commercial solutions include Data Harmony[42] and Synaptica[43], providing similar functionalities.

[34] We excluded the Library Classification for Environmental Science from our analysis because it was developed over five decades ago and is no longer actively maintained or used by the scientific community.

[35] German Centre for Higher Education Research and Science Studies (DZHW): www.dzhw.eu.

[36] Research Core Dataset: https://w3id.org/kdsf-ffk. Despite being a KOS of academic disciplines, RCD was not included in our analysis in Section 4 because it is available only in German (see selection criteria in Section 3.2).

[37] Trello: https://trello.com.

[38] Protégé: https://protege.stanford.edu.

[39] Semantic MediaWiki: https://www.semantic-mediawiki.org.

[40] VocBench: https://vocbench.uniroma2.it.

[41] PoolParty: https://www.poolparty.biz.

[42] Data Harmony: https://www.accessinn.com/data-harmony.

[43] Synaptica: https://www.synaptica.com.

In brief, a wide collection of tools facilitating the integration of the different KOSs is available. However, all these tools are general-purpose and built to support a multitude of use cases. As a result, users unfamiliar with the relevant technologies and the semantic web might find them quite complex and difficult to learn. We still lack user-friendly tools that are able to effectively support the creation and curation of KOSs.

5.3. Improving the Language Coverage

English is the *de facto* language of science nowadays (Sugimoto & Larivière, 2018). Most international conferences are held in English, and the world's top scientific journals are published in English. However, some countries, like China, Russia, and Japan, have several dedicated journals and conferences publishing scientific papers in their own language. It is crucial then to have KOSs in other languages as well so as to facilitate interoperability and the cross-language exploration and exploitation of digital artefacts (Lei Zeng & Mai Chan, 2004). This would also benefit students in non-English-speaking countries, allowing them to understand and navigate KOSs without requiring fluency in English.

In Section 4.3.6, we acknowledged that some KOSs already have partial translations in other languages, such as the *Dewey Decimal Classification*, the *Agrovoc Thesaurus*, and the *UNESCO Thesaurus*. However, the process is far from being complete. The *UNESCO Thesaurus* is available in only five languages (English, Arabic, Russian, French, and Spanish). The *Agrovoc Thesaurus* has a very wide variety of languages; however, not all the 42 languages are equally represented, as reported in Figure 6. Notably, the Food and Agriculture Organization, which coordinates *Agrovoc*, allows international institutions to contribute to *Agrovoc* by authoring translations. This is an interesting solution that may be beneficial for several other KOSs.

Overall, the support for languages different from English is still very poor for most of the KOSs analyzed in this survey. We need further work to produce resources able to support multiple languages (Tudhope & Nielsen, 2006). In this context, large language models offer a promising solution, as they have proved to be highly effective on automatic language translation (Lu, Zhu et al., 2024).

5.4. Reconciliating Expert Disagreements

One of the major challenges appearing when working in collaborative environments is conflict management, and the process of building and integrating KOSs is not immune to this. Indeed, Chilton, Little et al. (2013) argue that the main characteristic of taxonomies, and by extension KOSs, is that they are subjective. Experts have different backgrounds, based mainly on the paradigms they inhabit (Kuhn, 1962), and therefore, they often tend to disagree on the properties and the structure of the system. Such disagreement also has an impact on the frequency of updates because it requires time to be addressed, with a consequent delay in the new release.

In the literature, we can find different works studying disagreement among experts in the context of building KOSs. Gu, Hripcsak et al. (2007) audited the semantic types assigned to *UMLS* concepts with the support of four experts. In particular, the experts assessed whether the current semantic types assigned to the concepts were correct. In some cases, the disagreement prevailed even after multiple rounds and was eventually resolved with a subsequent discussion. This experience indicates that the process of mitigating disagreement between experts is quite challenging and certainly time consuming. Osborne, Muccini et al. (2019) analyzed the

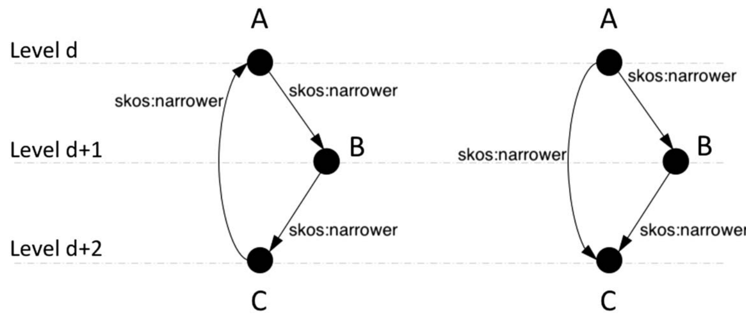


Figure 8. Example of inconsistencies within the KOSs. On the left the cyclic inconsistency and on the right the redundant inconsistency.

agreement of experts in characterizing articles according to research areas in Software Architecture and highlighted that, in this domain, most of the disagreement was between domain experts of different seniority (e.g., Full Professor vs. PhD students). Fan and Friedman (2007) also explored the issue of disagreement among experts on research topics. They developed a classifier to annotate research documents with concepts derived from *UMLS* and then asked experts to categorize *UMLS* concepts into nine different types. While the interannotator agreement was relatively high (0.82), their ablation study revealed that experts often disagreed on vague concepts. For instance, the concept of "promotion" was classified both as a behavior and as a biological function. The authors emphasize that unresolved disagreements can introduce errors into gold standards, which in turn can propagate into downstream applications.

In conclusion, creating and integrating different KOSs is a collaborative effort, leading to experts' disagreement. We thus need to develop new tools and methodologies for handling disagreement and reconciling different suggestions.

5.5. Assessing the Quality of KOSs

A high-quality KOS must be coherent at two different levels: *structural* and *conceptual* (Ayele, Chevallet et al., 2012; Morrey, Geller et al., 2009; Raad & Cruz, 2015).

At the structural level, the KOS can present logical inconsistencies, i.e., contradictions hindering the integrity of the system (Ayele et al., 2012). Here, we illustrate two of the most frequent cases of logical inconsistencies: cyclic and transitive. Figure 8 shows the two examples, and specifically the arrows identify the `skos:narrower`[44] relationship, for instance "A" → "B" equals <*A, skos:narrower, B*> meaning "B" is a narrower concept of "A," and hence "A" is broader than "B." The first case, shown on the left, is the *cyclic* inconsistency (also known as *loop*): "A" is broader than "B," "B" is broader than "C" and "C" is broader than "A." This is inconsistent because the concept *C* cannot be simultaneously narrower (via "B") and broader than "A." Mougin and Bodenreider (2005) highlight that automatically addressing this problem is challenging because identifying the appropriate connection to break is difficult, and performing this operation may introduce additional errors.

The second case of inconsistency is when there is a *transitive hierarchy*[45]: "A" is broader than "B," "B" is broader than "C," but "A" is also broader than "C." Although, in both cases "A" is broader than "C," hence it may not be considered as a real mistake; however, it adds

[44] `skos:narrower`: https://www.w3.org/TR/skos-reference/#semantic-relations.
[45] Transitive hierarchies: https://www.w3.org/TR/skos-primer/#sectransitivebroader.

confusion to the hierarchical structure which needs to be addressed. Besides, by convention, `skos:narrower` and `skos:broader` must be used to assert a direct or immediate hierarchical link between concepts and are not declared as transitive properties.

These two cases have been simplified just for the sake of clarity, but such inconsistencies might occur across longer chains of concepts (Bodenreider, 2001). However, ontology reasoners can help to identify these logical inconsistencies (Jiménez-Ruiz et al., 2011).

At a conceptual level, several challenges can hinder the quality of a KOS. These are related to the *accuracy* (Lambe, 2014), i.e., how the concepts are related to each other, such as ensuring correct hierarchical relationships between two subjects and addressing ambiguities in preserving multiple semantic views. Additional challenges are instead related to the *completeness* of the concepts (Lambe, 2014), such as whether an area needs to be refined or enriched with subareas; whether the description of concepts, including definition, scope, and editorial information, is appropriate; and if all related terms are correct and sufficient.

The literature presents various approaches for assessing the quality of KOSs (Morrey et al., 2009). For instance, Raad and Cruz (2015) discusses four categories of ontology evaluation approaches: gold standard based, corpus based, task based, and criteria based. Gold-standard-based evaluation approaches compare the KOS with a reference system. Corpus-based evaluation approaches assess the extent of coverage of the system in a given domain. Task-based evaluation approaches measure to what extent the system helps to improve a task. Finally, criteria-based evaluation approaches measure the extent of a system in adhering to certain criteria. The wide range of evaluation methods highlights the complexity of assessing the quality of a KOS. Indeed, for gold-standard-based evaluations, Raad and Cruz (2015) point out that it is hard to find a suitable gold standard because it needs to be created with the same conditions as the evaluated system. Corpus-based evaluation approaches also have similar issues. On the other hand, task-based and criteria-based evaluation approaches have fewer challenges to tackle, but they fall short in assessing KOSs at the conceptual level (i.e., accuracy and completeness).

In conclusion, when creating or integrating KOSs it is crucial to continuously assess the quality of the resulting knowledge representation (Shvaiko & Euzenat, 2008). The community needs to further advance tools, metrics, and methodologies in this space.

5.6. Handling Ambiguous Labels

Polysemy is a linguistic feature whereby a term has multiple meanings in different contexts. For instance, “Java” is both a programming language and an island in Indonesia. Polysemy has a negative effect during the process of building and aligning KOSs because it introduces ambiguity (Djeddi & Khadir, 2014; Johnson, Bretonnel Cohen, & Hunter, 2007).

In the literature, we can find several approaches attempting to disambiguate the various senses of a term. Bella, Giunchiglia, and McNeill (2017) mapped *Eurovoc* with *Universal Decimal Classification* (UDC)[46], and they observed that indeed polysemy induces the appearance of false positives, matching together terms that are syntactically similar but have different semantics. To address this problem, they created a Word Sense Disambiguation (WSD) method. Other methods developing WSD techniques are applied to *UMLS* (McInnes, Pedersen, & Carlis, 2007; Widdows, Peters et al., 2003); however, these methods necessitate

[46] We did not include UDC in our analysis because it has restricted access and requires paying a license.

the surrounding context of a term and struggle to infer its correct sense when applied directly to concept labels. In the context of ontologies, Pisanelli, Gangemi et al. (2004) suggest formally representing the several meanings of a word, specifying their types. For instance, the "Java" entity, in the sense of programming language, can be defined both as a `skos:Concept` as well as the type that defines it as a programming language. In this way, the two Java entities will be properly disambiguated through their specific type. Gu, Perl et al. (2004) audited *UMLS* and found several polysemic concepts, and in a shared view with Pisanelli et al. (2004), they also suggest replacing the polysemous concept with new different concepts according to the different intersecting semantic types.

In general, disambiguating subjects according to an explicit representation of their meaning is the *de facto* solution to tackle ambiguous labels. However, it can also lead to some issues, because the resulting KOSs are more complex (Shi, Maly et al., 2011), harder to maintain, and the type assignment may cause disagreements among domain experts (Fan & Friedman, 2007).

5.7. Automatic Generation and Updating KOSs

The current landscape of KOSs reveals gaps in coverage for certain areas (see Section 4.1) and highlights the vulnerability of some KOSs to becoming outdated (see Section 4.3.4). Given the critical role KOSs play in various downstream services, the scientific community should prioritize developing more automatic tools that facilitate their creation, maintenance, and continuous updating.

The literature presents various automated solutions for creating KOSs. These can either be employed to generate KOSs from scratch (e.g., Osborne & Motta, 2015), or to scan the recent literature and extend existing ones (e.g., Huang, Xie et al., 2020; Pisu, Pompianu et al., 2024). Klink-2 (Osborne & Motta, 2015) demonstrated high effectiveness in automatically generating large-scale and granular ontologies of research areas. On a similar note, Shang, Zhang et al. (2020) introduced NetTaxo, which is an automated topic taxonomy construction framework that leverages both text data and network structures to build hierarchical topic representations, whereas, Zhang, Tao et al. (2018) presented TaxoGen, an unsupervised method that employs term embeddings and hierarchical clustering to recursively build a topic taxonomy. On the other hand, Huang et al. (2020) introduced a technique for building topic taxonomies guided by seed concepts. Given a text corpus and an initial taxonomy of concepts, their approach develops a more comprehensive taxonomy. This approach is potentially valuable in the context of updating existing KOSs.

In recent years, we have witnessed the emergence of new automatic approaches that leverage advanced language models. For instance, Pisu et al. (2024) introduced an AI-driven pipeline that extracts research concepts from articles and then determines their semantic relationships (hierarchical or synonymous) using SciBERT (Beltagy, Lo et al., 2019). Other approaches rely on Large Language Models, which possess a deeper understanding of language, enabling them to identify semantic relationships between concepts more accurately (Aggarwal, Salatino et al., 2024; Revenko, Breit et al., 2024). This challenge is attracting growing attention from the community due to the new possibilities offered by AI.

5.8. Automatic Classification of Items

The main objective of KOSs is to categorize a vast amount of items, such as articles, books, courses, patents, software, and experiment materials. Classifying them manually may be unfeasible on a large scale. In recent years, various methods have been developed for the automatic classification of these items. The literature typically categorizes these approaches into two main types: supervised and unsupervised methods.

Supervised methods need to be trained according to a set of labeled samples, each associated with one or more research topics to learn from. The resulting model can be used to automatically classify new items according to the relevant research topics. For instance, Mai, Galke, and Scherp (2018) employed deep learning techniques to develop a classifier applied to a set of papers annotated with the *STW Thesaurus for Economics* and *MeSH*. Similarly, Chernyak (2015) presented a classifier in "Computer Science" with topics drawn from the *ACM Computing Classification System*. Caragea, Bulgarov, and Mihalcea (2015) trained their classifiers on a corpus of 3,186 papers distributed over six classes (subjects): *agents, artificial intelligence, information retrieval, human computer interaction, machine learning*, and *databases*. Kandimalla et al. (2021) developed a deep attentive neural network for classifying papers according to 104 (out of 254) *Web of Science subject categories*. Their classifier was trained on 9 million abstracts from Web of Science.

Recently, the OpenAlex team developed a large deep learning model that leverages the research paper's title, abstract, citations, and journal name to assign classifications drawn from the 4,500 topics within the *OpenAlex Topics* (OpenAlex, 2024). This model demonstrates modest accuracy, with the correct label appearing as the top prediction for 53% of papers and within the top 10 predictions for 73% of papers. To date, it represents a potentially unique example of a deep learning model deployed with such a large number of categories. Supervised approaches often face two significant limitations. First, they typically struggle with handling a large number of categories, which can lead to underperformance when applied to very large KOSs. Second, they are impractical for KOSs that lack a sufficiently large dataset of labeled items.

Unsupervised methods in this space typically aim to associate research areas described in the KOSs to specific text segments using semantic similarity metrics, word embeddings, and a variety of NLP techniques. Their main advantage is that they do not require a training set.

For instance, the CSO Classifier (Salatino et al., 2022) is a tool designed to classify research documents within the field of Computer Science. It processes the textual elements of a research document, such as the title, abstract, and keywords, and outputs the relevant topics based on the *Computer Science Ontology*. This classifier consists of two main modules. The first aims at identifying the *CSO* concepts that are explicitly mentioned in the document. The second module leverages word embeddings to infer semantically related topics. Finally, an additional component postprocesses the identified topics and removes outliers. Soldaini and Goharian (2016) developed QuickUMLS, a scalable approach for annotating documents with concepts drawn from *UMLS*. This approach takes the document, extracts parts of speech, then preprocesses the text to identify valid tokens and finally matches them against the labels in *UMLS*. It is able to get comparable results with state-of-the-art solutions, but it tends to confuse polysemic terms. MetaMap (Aronson & Lang, 2010) is a similar approach for annotating documents with *UMLS* concepts. MetaMap extracts tokens and their part of speech, then it selects the candidate and matches them against *UMLS* labels. It also implements word sense disambiguation to identify the right sense of the concept in the case of polysemic terms. However, it suffers from scalability issues and only identifies concepts whose labels are syntactically available in the text. Savova, Masanz et al. (2010) developed cTAKES, which is an approach for extracting *UMLS* concepts from medical records. It implements a named-entity recognition approach, identifying *UMLS* terms within a noun-phrase lookup window. However, cTAKES currently does not resolve ambiguities.

A well-known limitation of unsupervised approaches is that they rely on language models, which typically need to be trained on a specific field. Therefore, they are typically unable to work cross-discipline. The emergence of large language models has opened up exciting new possibilities due to their ability to generalize across diverse research areas (Kojima, Gu et al., 2023). However, further investigation is necessary before they can be confidently deployed in real-world applications.

In conclusion, only a few fields (e.g., "Medicine," "Biology," "Chemistry," "Computer Science," "Economics") currently have access to high-quality classifiers based on a granular representation of the discipline. Furthermore, these classifiers still face several technical limitations. As a result, it is essential to develop new solutions that not only expand the range of disciplines involved, but also improve accuracy and scalability.

6. THREATS TO VALIDITY

This section examines potential threats to validity of our study. We identify four key areas where the validity of our study could be challenged: internal validity, external validity, construct validity, and conclusion validity, as discussed in Wohlin, Runeson et al. (2024). In the following discussion, we evaluate these potential threats and detail the measures we have taken to minimizse their impact.

6.1. Internal Validity

Internal validity in surveys pertains to the rigor and accuracy of the adopted methodology. To guarantee the replicability of our survey, we carefully devised a methodologically sound protocol that included systematic and transparent phases for selecting knowledge organization systems. The initial protocol was developed by the first author and subsequently reviewed and refined by coauthors to establish consensus before initiating the survey process. We utilized well-known search engines (e.g., Google Scholar, Google) and various sources such as publishers, Wikipedia, and interconnections among the collected KOSs. Additionally, we consulted domain experts in relevant research fields to expand our result set further.

We performed a multistage selection process to ensure a rigorous evaluation and minimize selection bias. Initially, the first author analyzed and selected all tools based on their description. Subsequently, all authors collaboratively conducted a comprehensive review of the shortlisted KOSs. In cases where information was unclear or unavailable, the first author directly contacted the curators of the respective KOSs for clarification.

Although we employed a systematic approach, biases could still arise from subjective decisions made during the application of inclusion and exclusion criteria. To address this, we conducted collaborative reviews of the shortlisted KOSs, which helped to minimize the impact of individual biases on the selection process.

In summary, although another research team replicating this study might find minor differences in the specific KOSs, the rigorous and systematic methodology used, combined with the collaborative nature of the process, strongly supports the internal validity of our results.

6.2. External Validity

External validity refers to how well the results of this survey can be generalized and applied to other contexts and areas of study. To maximize the applicability of our findings, we utilized a variety of sources when choosing the KOSs for conducting this analysis. While we aimed for a comprehensive identification of tools, some relevant KOSs may have been inadvertently excluded due to limitations in the search engines or query terms used. This could occur if KOSs were not adequately described or indexed using appropriate keywords. To address this, we continuously refined our search terms and consulted domain experts to ensure a wider range of potential KOSs were captured. Additionally, we explored the external links of identified KOSs to discover further relevant tools.

With regard to the inclusion and exclusion criteria, we identified two main potential threats to external validity. The first concerns the exclusion of KOSs that do not offer an English version. This was set due to the predominance of English as the language of scientific communication (Sugimoto & Larivière, 2018). Additionally, the absence of English versions would have hindered our ability to conduct in-depth analyses, such as the one presented in Figure 3. The second threat arises from the exclusion of KOSs that are designed specifically for individual digital libraries and have not been widely adopted by the broader community. This exclusion was necessary to avoid an unmanageable increase in the number of KOS candidates, which would have made the analysis impractical. Additionally, these KOSs are often customized to fit the unique content of their respective libraries, which could result in a skewed portrayal of the broader scientific landscape.

6.3. Construct Validity

Construct validity refers to the extent to which the operational measures used in a study accurately represent the concepts under investigation. In our survey, the primary concern is whether the 15 analyzed features comprehensively cover all relevant aspects.

To address any potential omissions in our analysis, all authors collaboratively and iteratively defined the five feature categories for evaluation (i.e., scope, structure, curation, external links, and usage) and the 15 specific features.

We recognize that our analysis may not have fully addressed all relevant aspects. For instance, evaluating the quality of each KOS could offer valuable insights. However, there is no universally accepted definition of "quality" in the context of KOSs, which could lead to potential bias if a specific definition were adopted. Moreover, such an evaluation would be time-consuming, expensive, and require specialized expertise across various scientific disciplines.

6.4. Conclusion Validity

Conclusion validity in surveys refers to how well the conclusions drawn are supported by the evidence and are reproducible. In our analysis, we placed great emphasis on minimizing threats to conclusion validity by using a systematic approach to identify relevant KOSs and extract the relevant features.

To ensure precise and reliable data collection, we developed a data extraction form based on the 15 selected features and a protocol for identifying relevant information.

Each author independently analyzed a sample of KOSs using this standardized form and protocol. Furthermore, for KOSs available in machine-readable formats, we created custom Python scripts to extract structural features and external links (available on our GitHub repository: https://github.com/angelosalatino/kos-rf). Our protocol also included cross-checking our analyses to guarantee accuracy and consistency.

A continuous challenge to the validity of our conclusions is the dynamic nature of KOSs. Many of them are updated annually, acquiring new concepts. Consequently, it is anticipated that many KOSs will evolve in the near future. While our findings provide a snapshot of the current landscape, they may not fully capture the ongoing developments in this field.

7. CONCLUSIONS

Knowledge organization systems of academic fields (e.g., term lists, taxonomies, thesauri, ontologies) are an important part of the academic ecosystem and enable the categorization, management, and retrieval of items and information. These solutions have become particularly important in the last few years given the ever-growing number of publications, the rise of Open Science, and the emergence of vast online repositories of articles, courses, and other academic materials (Auer et al., 2018).

This article provides a systematic overviewof 45 KOSs, with 23 focusing on a single field and 22 covering multiple fields. We propose an analysis framework that characterizes them according to five main dimensions: scope, structure, curation, usage, and links to other KOSs. The comparative table describing the 45 KOSs according to the 15 features is available at https://doi.org/10.48366/R732033, as a living review. All the code produced and the data retrieved during this analysis are openly available at https://github.com/angelosalatino/kos-rf. These resources can be freely reused by researchers to rerun the analysis in the future.

Our findings indicate that the current generation of KOSs requires substantial enhancements in scope, quality, and granularity. Notably, seven disciplines ("History," "Political Science," "Environmental Science," "Material Science," "Geography," "Sociology," and "Business") lack dedicated, field-specific KOSs. Additionally, among the existing multidisciplinary KOSs, only five provide comprehensive coverage of the 19 fields analyzed in this study.

The analyzed KOSs exhibit considerable diversity in both the number of concepts and structural depth. Some KOSs include as few as 48 concepts (e.g., *Fields of research and development*) with a relatively shallow structure, while others encompass more than 3 million concepts and feature a very deep structure, extending beyond 30 levels (e.g., *Open Biological and Biomedical Ontology* and *Unified Medical Language System*). The majority of the KOSs analyzed (24) are still traditional taxonomies. However, a growing number of more recent KOSs (18) are now based on ontologies. Regarding hierarchical structures, there is a fairly balanced distribution: 24 KOSs utilize a poly-hierarchical structure, while 20 employ a mono-hierarchical one. Traditionally, KOSs have been created and maintained manually. However, in recent years, there has been a significant shift toward more automated methods. Consequently, eight KOSs, including *OpenAlex Topics* and *Microsoft's Fields of Study*, have implemented modern automatic or semiautomatic pipelines for their generation and updating processes. Eighteen KOSs are updated annually, demonstrating a significant commitment to maintaining up-to-date resources in some disciplines (e.g., "Medicine," "Engineering," "Agriculture," and "Computer Science"). Finally, 26 KOSs are available under open licenses.

Our findings indicate that there is currently no existing multifield KOS that simultaneously is comprehensive in topic coverage, granular, consistently updated, and openly accessible (see Figure 7). The creation of such a KOS could revolutionize the field by dramatically improving content organization across digital libraries, simplifying the process of research information gathering, facilitating the monitoring of research and development activities, ensuring high data quality, and supporting evidence-based policymaking.

We have identified key future research directions in this domain, along with the challenges that accompany them. These priorities include generating higher-quality multifield KOSs, developing new automatic methods for integrating and updating KOSs, adopting standardized formats, expanding language coverage, and evaluating KOSs across diverse characteristics. Furthermore, there is a critical need for more precise and scalable methods for automatically classifying articles according to these systems.

Overall, this is a promising area of research that has yet to fully capitalize on recent advances in generative AI, which hold significant potential to drive progress in the field. To achieve this, we believe it is essential for the Open Science, Digital Libraries, and Artificial Intelligence communities to collaborate in developing a unified framework of interlinked resources. We hope that this survey paper serves as a meaningful first step toward this direction.

ACKNOWLEDGMENTS

The authors wish to thank the reviewers for their constructive feedback, which is reflected in this version of the manuscript. They also gratefully acknowledge the Open University Library for providing access to the Dewey Decimal Classification.

AUTHOR CONTRIBUTIONS

Angelo Salatino: Conceptualization, Formal analysis, Methodology, Software, Visualization, Writing—original draft. Tanay Aggarwal: Formal analysis, Software. Andrea Mannocci: Conceptualization, Formal analysis, Methodology, Writing—review & editing. Francesco Osborne: Conceptualization, Formal analysis, Methodology, Writing—review & editing. Enrico Motta: Formal analysis, Writing—review & editing.

COMPETING INTERESTS

The authors have no competing interests.

FUNDING INFORMATION

No funding was received for this research.

DATA AVAILABILITY

The code for processing the analyzed KOSs is available on a GitHub repository: https://github.com/angelosalatino/kos-rf. A complete table of all KOSs and their analyzed features can be found at https://doi.org/10.48366/R732033.

REFERENCES

Aggarwal, T., Salatino, A., Osborne, F., & Motta, E. (2024). Identifying semantic relationships between research topics using large language models in a zero-shot learning setting. In A. Salatino, A. Mannocci, F. Osborne, S. Schimmler, & G. Rehm (Eds.), *Proceedings of the 4th International Workshop on Scientific Knowledge: Representation, Discovery, and Assessment*. https://ceur-ws.org/Vol-3780/paper3.pdf

Angioni, S., Salatino, A., Osborne, F., Birukou, A., Recupero, D. R., & Motta, E. (2022a). Leveraging knowledge graph technologies to assess journals and conferences at Springer Nature. In U. Sattler et al. (Eds.), *The Semantic Web – ISWC 2022* (pp. 735–752). Cham: Springer. https://doi.org/10.1007/978-3-031-19433-7_42

Angioni, S., Salatino, A., Osborne, F., Recupero, D. R., & Motta, E. (2022b). AIDA: A knowledge graph about research dynamics in academia and industry. *Quantitative Science Studies*, *2*(4), 1356–1398. https://doi.org/10.1162/qss_a_00162

Aronson, A. R., & Lang, F.-M. (2010). An overview of MetaMap: Historical perspective and recent advances. *Journal of the American Medical Informatics Association*, *17*(3), 229–236. https://doi.org/10.1136/jamia.2009.002733, PubMed: 20442139

Auer, S., Kovtun, V., Prinz, M., Kasprzik, A., Stocker, M., & Vidal, M. E. (2018). Towards a knowledge graph for science. In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics (WIMS '18)*. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/3227609.3227689

Ayala, O., & Bechard, P. (2024). Reducing hallucination in structured outputs via Retrieval-Augmented Generation. In Y. Yang, A. Davani, A. Sil, & A. Kumar (Eds.), *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 6: Industry Track)* (pp. 228–238). Mexico City, Mexico:

Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.naacl-industry.19

Ayele, D., Chevallet, J.-P., Kassie, G., & Meshesha, M. (2012). Enhancing semantic relation quality of UMLS knowledge sources. In *Proceedings of the International Conference on Management of Emergent Digital Ecosystems (MEDES '12)* (pp. 59–66). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2457276.2457289

Bella, G., Giunchiglia, F., & McNeill, F. (2017). Language and domain aware lightweight ontology matching. *Journal of Web Semantics*, *43*, 1–17. https://doi.org/10.1016/j.websem.2017.03.003

Beltagy, I., Lo, K., & Cohan, A. (2019). SciBERT: A pretrained language model for scientific text. In K. Inui, J. Jiang, V. Ng, & X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3615–3620). Hong Kong: Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1371

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, *284*(5), 34–43. https://doi.org/10.1038/scientificamerican0501-34

Bodenreider, O. (2001). Circular hierarchical relationships in the UMLS: Etiology, diagnosis, treatment, complications and prevention. *Proceedings of the AMIA Symposium*, *2001*, 57–61. PubMed: 11825155

Bodenreider, O. (2004). The unified medical language system (UMLS): Integrating biomedical terminology. *Nucleic Acids Research*, *32*(suppl_1), D267–D270. https://doi.org/10.1093/nar/gkh061, PubMed: 14681409

Bolaños, F., Salatino, A., Osborne, F., & Motta, E. (2024). Artificial intelligence for literature reviews: Opportunities and challenges. *Artificial Intelligence Review*, *57*, 259. https://doi.org/10.1007/s10462-024-10902-3

Borges, M. V. M., & dos Reis, J. C. (2019). Semantic-enhanced recommendation of video lectures. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)* (Vol. 2161, pp. 42–46). https://doi.org/10.1109/ICALT.2019.00013

Brainard, J. (2020). Scientists are drowning in COVID-19 papers. Can new tools keep them afloat? *Science*, May 13. https://doi.org/10.1126/science.abc7839

Caragea, C., Bulgarov, F., & Mihalcea, R. (2015). Co-training for topic classification of scholarly data. In L. Màrquez, C. Callison-Burch, & J. Su (Eds.), *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing* (pp. 2357–2366). Lisbon, Portugal: Association for Computational Linguistics. https://doi.org/10.18653/v1/D15-1283

Chatzopoulos, S., Vergoulis, T., Kanellos, I., Dalamagas, T., & Tryfonopoulos, C. (2020). ArtSim: Improved estimation of current impact for recent articles. In L. Bellatreche et al. (Eds.), *ADBIS, TPDL and EDA 2020 Common Workshops and Doctoral Consortium* (pp. 323–334). Cham: Springer. https://doi.org/10.1007/978-3-030-55814-7_27

Chernyak, E. (2015). An approach to the problem of annotation of research publications. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining* (pp. 429–434). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2684822.2697032

Chilton, L. B., Little, G., Edge, D., Weld, D. S., & Landay, J. A. (2013). Cascade: Crowdsourcing taxonomy creation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1999–2008). https://doi.org/10.1145/2470654.2466265

Chowdhury, G. (2010). From digital libraries to digital preservation research: The importance of users and context. *Journal of Documentation*, *66*(2), 207–223. https://doi.org/10.1108/00220411011023625

Cleverley, P. H., & Burnett, S. (2015). The best of both worlds: Highlighting the synergies of combining manual and automatic knowledge organization methods to improve information search and discovery. *Knowledge Organization*, *42*(6), 428–444. https://doi.org/10.5771/0943-7444-2015-6-428

Dai, S., You, R., Lu, Z., Huang, X., Mamitsuka, H., & Zhu, S. (2020). FullMeSH: Improving large-scale MeSH indexing with full text. *Bioinformatics*, *36*(5), 1533–1541. https://doi.org/10.1093/bioinformatics/btz756, PubMed: 31596475

Daiber, J., Jakob, M., Hokamp, C., & Mendes, P. N. (2013). Improving efficiency and accuracy in multilingual entity extraction. In *Proceedings of the 9th International Conference on Semantic Systems* (pp. 121–124). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2506182.2506198

Declerck, T. (2013). Integration of the thesaurus for the social sciences (TheSoz) in an information extraction system. In P. Lendvai & K. Zervanou (Eds.), *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 90–95). Sofia, Bulgaria: Association for Computational Linguistics. https://aclanthology.org/W13-2712/

Dessì, D., Osborne, F., Reforgiato Recupero, D., Buscaldi, D., & Motta, E. (2022). CS-KG: A large-scale knowledge graph of research entities and claims in computer science. In *The Semantic Web – ISWC 2022* (pp. 678–696). Cham: Springer. https://doi.org/10.1007/978-3-031-19433-7_39

Dhombres, F., & Bodenreider, O. (2016). Interoperability between phenotypes in research and healthcare terminologies—Investigating partial mappings between HPO and SNOMED CT. *Journal of Biomedical Semantics*, *7*, 3. https://doi.org/10.1186/s13326-016-0047-3, PubMed: 26865946

Dimou, A., Vander Sande, M., Colpaert, P., Verborgh, R., Mannens, E., & Van de Walle, R. (2014). RML: A generic language for integrated RDF mappings of heterogeneous data. In C. Bizer, T. Heath, S. Auer, & T. Berners-Lee (Eds.), *Proceedings of the 7th Workshop on Linked Data on the Web* (Vol. 1184). https://ceur-ws.org/Vol-1184/ldow2014_paper_01.pdf

Ding, Y., Rousseau, R., & Wolfram, D. (Eds.). (2014). *Measuring scholarly impact: Methods and practice*. Cham: Springer. https://doi.org/10.1007/978-3-319-10377-8

Djeddi, W. E., & Khadir, M. T. (2014). A novel approach using context-based measure for matching large scale ontologies. In L. Bellatreche & M. K. Mohania (Eds.), *Data warehousing and knowledge discovery* (pp. 320–331). Cham: Springer. https://doi.org/10.1007/978-3-319-10160-6_29

Dunne, E., & Hulek, K. (2020). Mathematics subject classification 2020. *Notices of the American Mathematical Society*, *67*(3), 410–411. https://doi.org/10.1090/noti2052

Erdogan, H., Erdem, E., & Bodenreider, O. (2010). Exploiting UMLS semantics for checking semantic consistency among UMLS concepts. *Studies in Health Technology and Informatics*, *160*(1), 749–753. https://doi.org/10.3233/978-1-60750-588-4-749, PubMed: 20841786

Fan, J.-W., & Friedman, C. (2007). Combining contextual and lexical features to classify UMLS concepts. *AMIA Annual Symposium Proceedings*, *2007*, 231–235. PubMed: 18693832

Genesereth, M. R., & Nilsson, N. J. (2012). *Logical foundations of artificial intelligence*. Morgan Kaufmann. https://doi.org/10.1016/C2009-0-27551-9 (Original work published 1987)

Gnoli, C., Golub, K., Haynes, D., Salaba, A., Shiri, A., & Slavic, A. (2024). Library catalog's search interface: Making the most of subject metadata. *Knowledge Organization*, *51*(3), 169–186. https://doi.org/10.5771/0943-7444-2024-3-169

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220. https://doi.org/10.1006/knac.1993.1008

Gu, H. H., Hripcsak, G., Chen, Y., Morrey, C. P., Elhanan, G., … Perl, Y. (2007). Evaluation of a UMLS auditing process of semantic type assignments. *AMIA Annual Symposium Proceedings*, *2007*, 294–298. PubMed: 18693845

Gu, H. H., Perl, Y., Elhanan, G., Min, H., Zhang, L., & Peng, Y. (2004). Auditing concept categorizations in the UMLS. *Artificial Intelligence in Medicine*, *31*(1), 29–44. https://doi.org/10.1016/j.artmed.2004.02.002, PubMed: 15182845

Halper, M., Morrey, C. P., Chen, Y., Elhanan, G., Hripcsak, G., & Perl, Y. (2011). Auditing hierarchical cycles to locate other inconsistencies in the UMLS. *AMIA Annual Symposium Proceedings*, *2011*, 529–536. PubMed: 22195107

Han, K., Yang, P., Mishra, S., & Diesner, J. (2020). WikiCSSH: Extracting computer science subject headings from Wikipedia. In L. Bellatreche et al. (Eds.), *ADBIS, TPDL and EDA 2020 Common Workshops and Doctoral Consortium* (pp. 207–218). Cham: Springer. https://doi.org/10.1007/978-3-030-55814-7_17

Hedden, H. (2010). Taxonomies and controlled vocabularies best practices for metadata. *Journal of Digital Asset Management*, *6*(5), 279–284. https://doi.org/10.1057/dam.2010.29

Hodge, G. M. (2000). *Systems of knowledge organization for digital libraries: Beyond traditional authority files* (No. 91). Digital Library Federation.

Huang, J., Xie, Y., Meng, Y., Zhang, Y., & Han, J. (2020). CoRel: Seed-guided topical taxonomy construction by concept learning and relation transferring. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 1928–1936). https://doi.org/10.1145/3394486.3403244

Ilgisonis, E. V., Pyatnitskiy, M. A., Tarbeeva, S. N., Aldushin, A. A., & Ponomarenko, E. A. (2022). How to catch trends using MeSH terms analysis? *Scientometrics*, *127*(4), 1953–1967. https://doi.org/10.1007/s11192-022-04292-y, PubMed: 35221395

International Organization for Standardization. (2013). *Information and documentation—Thesauri and interoperability with other vocabularies. Part 2: Interoperability with other vocabularies* (Standard). https://www.iso.org/standard/53658.html

Jiménez-Ruiz, E., Grau, B. C., Horrocks, I., & Berlanga, R. (2011). Logic-based assessment of the compatibility of UMLS ontology sources. *Journal of Biomedical Semantics*, *2*(Suppl. 1), S2. https://doi.org/10.1186/2041-1480-2-S1-S2, PubMed: 21388571

Jing, X. (2021). The Unified Medical Language System at 30 years and how it is used and published: Systematic review and content analysis. *JMIR Medical Informatics*, *9*(8), e20675. https://doi.org/10.2196/20675, PubMed: 34236337

Johnson, H. L., Bretonnel Cohen, K., & Hunter, L. (2007). A fault model for ontology mapping, alignment, and linking systems. In *Pacific Symposium on Biocomputing 2007* (pp. 233–244). World Scientific. PubMed: 17990495

Kalfoglou, Y., & Schorlemmer, M. (2003). Ontology mapping: The state of the art. *The Knowledge Engineering Review*, *18*(1), 1–31. https://doi.org/10.1017/S0269888903000651

Kandimalla, B., Rohatgi, S., Wu, J., & Giles, C. L. (2021). Large scale subject category classification of scholarly papers with deep attentive neural networks. *Frontiers in Research Metrics and Analytics*, *5*, 600382. https://doi.org/10.3389/frma.2020.600382, PubMed: 33870061

Kang, Y.-B., Li, Y.-F., & Coppel, R. L. (2015). Capturing researcher expertise through MeSH classification. In *Proceedings of the 8th International Conference on Knowledge Capture*, Article 6. New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2815833.2815837

Kendall, E. F., & McGuinness, D. L. (2019). *Ontology engineering*. Cham: Springer. https://doi.org/10.1007/978-3-031-79486-5

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2023). Large language models are zero-shot reasoners. *arXiv*. https://doi.org/10.48550/arXiv.2205.11916

Kopácsi, S., Hudak, R., & Ganguly, R. (2017). Implementation of a classification server to support metadata organization for long term preservation systems. *Communications of the Association of Austrian Librarians*, *70*(2), 225–243. https://doi.org/10.31263/voebm.v70i2.1897

Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Lambe, P. (2014). *Organising knowledge: Taxonomies, knowledge and organisational effectiveness*. Elsevier. https://www.sciencedirect.com/book/9781843342274/organising-knowledge#book-info

Lei Zeng, M., & Mai Chan, L. (2004). Trends and issues in establishing interoperability among knowledge organization systems. *Journal of the American Society for Information Science and Technology*, *55*(5), 377–395. https://doi.org/10.1002/asi.10387

Löffler, F., Wesp, V., Babalou, S., Kahn, P., Lachmann, R., … König-Ries, B. (2020). ScholarLensViz: A visualization framework for transparency in semantic user profiles. In K. Taylor, R. Gonçalves, F. Lecue, & J. Yan (Eds.), *Proceedings of the ISWC 2020 Demos and Industry Tracks*. https://ceur-ws.org/Vol-2721/paper485.pdf

Lu, Y., Zhu, W., Li, L., Qiao, Y., & Yuan, F. (2024). LLaMAX: Scaling linguistic horizons of LLM by enhancing translation capabilities beyond 100 languages. In Y. Al-Onaizan, M. Bansal, & Y.-N. Chen (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2024* (pp. 10748–10772). Miami, FL: Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.findings-emnlp.631

Mai, F., Galke, L., & Scherp, A. (2018). Using deep learning for title-based semantic subject indexing to reach competitive performance to full-text. In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries* (pp. 169–178). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/3197026.3197039

Martínez-González, M. M., & Alvite-Díez, M.-L. (2019). Thesauri and semantic web: Discussion of the evolution of thesauri toward their integration with the semantic web. *IEEE Access*, *7*, 153151–153170. https://doi.org/10.1109/ACCESS.2019.2948028

Mazzocchi, F. (2018). Knowledge organization system (KOS): An introductory critical account. *Knowledge Organization*, *45*(1), 54–78. https://doi.org/10.5771/0943-7444-2018-1-54

McInnes, B. T., Pedersen, T., & Carlis, J. (2007). Using UMLS concept unique identifiers (CUIs) for word sense disambiguation in the biomedical domain. *AMIA Annual Symposium Proceedings*, *2007*, 533–537. PubMed: 18693893

Mitchell, J. S. (2001). Relationships in the Dewey Decimal Classification system. In C. A. Bean & R. Green (Eds.), *Relationships in the organization of knowledge* (pp. 211–226). Dordrecht: Springer. https://doi.org/10.1007/978-94-015-9696-1_14

Montiel-Ponsoda, E., De Cea, G. A., Gómez-Pérez, A., & Peters, W. (2011). Enriching ontologies with multilingual information. *Natural Language Engineering*, *17*(3), 283–309. https://doi.org/10.1017/S1351324910000082

Morrey, C. P., Geller, J., Halper, M., & Perl, Y. (2009). The neighborhood auditing tool: A hybrid interface for auditing the UMLS. *Journal of Biomedical Informatics*, *42*(3), 468–489. https://doi.org/10.1016/j.jbi.2009.01.006, PubMed: 19475725

Motta, E. (1999). *Reusable components for knowledge modelling: Case studies in parametric design problem solving*. Amsterdam: IOS Press. https://books.google.co.uk/books?id=S6g8FTcQAVYC

Mougin, F., & Bodenreider, O. (2005). Approaches to eliminating cycles in the UMLS metathesaurus: Naïve vs. formal. *AMIA Annual Symposium Proceedings*, *2005*, 550–554. PubMed: 16779100

Mu, X., Lu, K., & Ryu, H. (2014). Explicitly integrating MeSH thesaurus help into health information retrieval systems: An empirical user study. *Information Processing & Management*, *50*(1), 24–40. https://doi.org/10.1016/j.ipm.2013.03.005

Musen, M. A. (2015). The Protégé project: A look back and a look forward. *AI Matters*, *1*(4), 4–12. https://doi.org/10.1145/2757001.2757003, PubMed: 27239556

National Information Standards Organization. (2010). *Guidelines for the construction, format, and management of monolingual controlled vocabularies* (Standard). https://doi.org/10.3789/ansi.niso.z39.19-2005R2010

Newman, D., Noh, Y., Talley, E., Karimi, S., & Baldwin, T. (2010). Evaluating topic models for digital libraries. In *Proceedings of the 10th Annual Joint Conference on Digital Libraries* (pp. 215–224). https://doi.org/10.1145/1816123.1816156

OpenAlex. (2024). *OpenAlex: End-to-end process for topic classification*. https://docs.google.com/document/d/1bDopkhuGieQ4F8gGNj7sEc8WSE8mvLZS/edit

Osborne, F., & Motta, E. (2015). Klink-2: Integrating multiple web sources to generate semantic topic networks. In M. Arenas et al. (Eds.), *The Semantic Web—ISWC 2015* (pp. 408–424). Cham: Springer. https://doi.org/10.1007/978-3-319-25007-6_24

Osborne, F., Motta, E., & Mulholland, P. (2013). Exploring scholarly data with Rexplore. In H. Alani et al. (Eds.), *The Semantic Web—ISWC 2013* (pp. 460–477). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-41335-3_29

Osborne, F., Muccini, H., Lago, P., & Motta, E. (2019). Reducing the effort for systematic reviews in software engineering. *Data Science*, *2*(1–2), 311–340. https://doi.org/10.3233/DS-190019

Ovalle-Perandones, M.-A., Gorraiz, J., Wieland, M., Gumpenberger, C., & Olmeda-Gómez, C. (2013). The influence of European Framework Programmes on scientific collaboration in nanotechnology. *Scientometrics*, *97*(1), 59–74. https://doi.org/10.1007/s11192-013-1028-2

Peng, C., Xia, F., Naseriparsa, M., & Osborne, F. (2023). Knowledge graphs: Opportunities and challenges. *Artificial Intelligence Review*, *56*(11), 13071–13102. https://doi.org/10.1007/s10462-023-10465-9, PubMed: 37362886

Pisanelli, D. M., Gangemi, A., Battaglia, M., & Catenacci, C. (2004). Coping with medical polysemy in the semantic web: The role of ontologies. In *Proceedings of the 11th World Congress on Medical Informatics* (pp. 416–419). https://doi.org/10.3233/978-1-60750-949-3-416

Pisu, A., Pompianu, L., Salatino, A., Osborne, F., Riboni, D., … Reforgiato Recupero, D. (2024). Leveraging language models for generating ontologies of research topics. In *Text2KG 2024: International Workshop on Knowledge Graph Generation from Text*. https://ceur-ws.org/Vol-3747/text2kg_paper6.pdf

Plate, K. H. (1966). Library classification for environmental science. *Journal (Water Pollution Control Federation)*, *38*(4), 580–584. https://www.jstor.org/stable/25035533, PubMed: 5929843

Pranckutė, R. (2021). Web of Science (WoS) and Scopus: The titans of bibliographic information in today's academic world. *Publications*, *9*(1), 12. https://doi.org/10.3390/publications9010012

Qiu, J., Zhao, R., Yang, S., & Dong, K. (2017). *Informetrics: Theory, methods and applications*. Singapore: Springer. https://doi.org/10.1007/978-981-10-4032-0

Raad, J., & Cruz, C. (2015). A survey on ontology evaluation methods. In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management* (pp. 179–186). SciTePress. https://doi.org/10.5220/0005591001790186

Rasch, R. F. (1987). The nature of taxonomy. *Image: Journal of Nursing Scholarship*, *19*(3), 147–149. https://doi.org/10.1111/j.1547-5069.1987.tb00613.x

Revenko, A., Breit, A., Mahmoud, S., Szabo, M., & Knap, T. (2024). Investigate the impact of contextual information on LLMs for taxonomy expansion. In A. Salatino, M. Alam, F. Ongenae, S. Vahdati, A.-L. Gentile, T. Pellegrini, & S. Jiang (Eds.), *Knowledge graphs in the age of language models and neuro-symbolic AI* (pp. 37–51). Amsterdam: IOS Press. https://doi.org/10.3233/SSW240005

Reymond, D. (2020). Patents information for humanities research: Could there be something? *Iberoamerican Journal of Science Measurement and Communication*, *1*(1), 6. https://doi.org/10.47909/ijsmc.02

Salatino, A., Angioni, S., Osborne, F., Recupero, D. R., & Motta, E. (2023). Diversity of expertise is key to scientific impact: A large-scale analysis in the field of computer science. In *27th International Conference on Science, Technology and Innovation Indicators (STI 2023)*. https://doi.org/10.55835/6442f3fd947802668eee976c

Salatino, A., Osborne, F., & Motta, E. (2022). CSO Classifier 3.0: A scalable unsupervised method for classifying documents in terms of research topics. *International Journal on Digital Libraries*, *23*(1), 91–110. https://doi.org/10.1007/s00799-021-00305-y

Salatino, A. A. (2019). Early detection of research trends. *PhD Thesis*. https://doi.org/10.21954/ou.ro.00010698

Salatino, A. A., Osborne, F., Birukou, A., & Motta, E. (2019). Improving editorial workflow and metadata quality at Springer Nature. In C. Ghidini, O. Hartig, M. Maleshkova, V. Svátek, I. Cruz, A. Hogan, J. Song, M. Lefrançois, & F. Gandon (Eds.), *The Semantic Web—ISWC 2019* (pp. 507–525). Cham: Springer. https://doi.org/10.1007/978-3-030-30796-7_31

Salatino, A. A., Osborne, F., & Motta, E. (2018). AUGUR: Forecasting the emergence of new research topics. In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries* (pp. 303–312). https://doi.org/10.1145/3197026.3197052

Salatino, A. A., Osborne, F., & Motta, E. (2020). Ontology extraction and usage in the scholarly knowledge domain. In G. Cota, M. Daquino, & G. L. Pozzato (Eds.), *Applications and practices in ontology design, extraction, and reasoning* (pp. 91–106). Amsterdam: IOS Press. https://doi.org/10.3233/SSW200037

Salatino, A. A., Thanapalasingam, T., Mannocci, A., Birukou, A., Osborne, F., & Motta, E. (2020). The Computer Science Ontology: A comprehensive automatically-generated taxonomy of research areas. *Data Intelligence*, *2*(3), 379–416. https://doi.org/10.1162/dint_a_00055

Salatino, A. A., Thanapalasingam, T., Mannocci, A., Osborne, F., & Motta, E. (2018). The Computer Science Ontology: A large-scale taxonomy of research areas. In D. Vrandečić et al. (Eds.), *The*

Semantic Web—ISWC 2018 (pp. 187–205). Cham: Springer. https://doi.org/10.1007/978-3-030-00668-6_12

Salvadores, M., Alexander, P. R., Musen, M. A., & Noy, N. F. (2013). BioPortal as a dataset of linked biomedical ontologies and terminologies in RDF. *Semantic Web*, *4*(3), 277–284. https://doi.org/10.3233/SW-2012-0086, PubMed: 25214827

Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., ... Chute, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. *Journal of the American Medical Informatics Association*, *17*(5), 507–513. https://doi.org/10.1136/jamia.2009.001560, PubMed: 20819853

Schandl, T., & Blumauer, A. (2010). PoolParty: SKOS thesaurus management utilizing linked data. In L. Aroyo et al. (Eds.), *The semantic web: Research and applications* (pp. 421–425). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-13489-0_36

Scharnhorst, A., Börner, K., & van den Besselaar, P. (Eds.). (2012). *Models of science dynamics: Encounters between complexity theory and information sciences*. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-23068-4

Shang, J., Zhang, X., Liu, L., Li, S., & Han, J. (2020). NetTaxo: Automated topic taxonomy construction from text-rich network. In *Proceedings of the Web Conference 2020* (pp. 1908–1919). https://doi.org/10.1145/3366423.3380259

Shi, H., Maly, K., Zeil, S., & Zubair, M. (2011). Comparison of ontology reasoning systems using custom rules. In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics* (Article 16). https://doi.org/10.1145/1988688.1988708

Shiri, A. A., Revie, C., & Chowdhury, G. (2002). Thesaurus-assisted search term selection and query expansion: A review of user-centred studies. *Knowledge Organization*, *29*(1), 1–19. https://doi.org/10.5771/0943-7444-2002-1-1

Shvaiko, P., & Euzenat, J. (2008). Ten challenges for ontology matching. In R. Meersman & Z. Tari (Eds.), *On the move to meaningful internet systems: OTM 2008* (pp. 1164–1182). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-88873-4_18

Sinha, A., Shen, Z., Song, Y., Ma, H., Eide, D., ... Wang, K. (2015). An overview of Microsoft Academic Service (MAS) and applications. In *Proceedings of the 24th International Conference on World Wide Web* (pp. 243–246). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/2740908.2742839

Sjögårde, P., & Ahlgren, P. (2020). Granularity of algorithmically constructed publication-level classifications of research publications: Identification of specialties. *Quantitative Science Studies*, *1*(1), 207–238. https://doi.org/10.1162/qss_a_00004

Sjögårde, P., & Didegah, F. (2022). The association between topic growth and citation impact of research publications. *Scientometrics*, *127*(4), 1903–1921. https://doi.org/10.1007/s11192-022-04293-x

Slater, L. T., Gkoutos, G. V., & Hoehndorf, R. (2020). Towards semantic interoperability: Finding and repairing hidden contradictions in biomedical ontologies. *BMC Medical Informatics and Decision Making*, *20*(Suppl. 10), 311. https://doi.org/10.1186/s12911-020-01336-2, PubMed: 33319712

Soldaini, L., & Goharian, N. (2016). QuickUMLS: A fast, unsupervised approach for medical concept extraction. In *Proceedings of Medical Information Retrieval (MedIR) Workshop at SIGIR* (pp. 1–4).

Solimando, A., Jiménez-Ruiz, E., & Guerrini, G. (2014). Detecting and correcting conservativity principle violations in ontology-to-ontology mappings. In P. Mika et al. (Eds.), *The Semantic Web—ISWC 2014* (pp. 1–16). Cham: Springer. https://doi.org/10.1007/978-3-319-11915-1_1

Stellato, A., Fiorelli, M., Turbati, A., Lorenzetti, T., van Gemert, W., ... Keizer, J. (2020). VocBench 3: A collaborative semantic web editor for ontologies, thesauri and lexicons. *Semantic Web*, *11*(5), 855–881. https://doi.org/10.3233/SW-200370

Stiller, J., Trkulja, V., Biesenbender, S., & Petras, V. (2021). *Entwicklung einer klassifikation für interdisziplinäre forschungsfelder im rahmen des kerndatensatz forschung: Dokumentation des projekts und der projektergebnisse* (Technical Report). Deutsches Zentrum für Hochschul- und Wissenschaftsforschung. https://web.archive.org/web/20211004184211/https://www.kerndatensatz-forschung.de/docs_ff/2021-04-30_ffk-projekt_dokumentation.pdf

Sugimoto, C. R., & Larivière, V. (2018). *Measuring research: What everyone needs to know*. New York, NY: Oxford University Press. https://doi.org/10.1093/wentk/9780190640118.001.0001

Thanapalasingam, T., Osborne, F., Birukou, A., & Motta, E. (2018). Ontology-based recommendation of editorial products. In D. Vrandečić et al. (Eds.), *The Semantic Web—ISWC 2018* (pp. 341–358). Cham: Springer. https://doi.org/10.1007/978-3-030-00668-6_21

Tudhope, D., & Nielsen, M. L. (2006). Introduction to knowledge organization systems and services. *New Review of Hypermedia and Multimedia*, *12*(1), 3–9. https://doi.org/10.1080/13614560600856433

Tudorache, T., Nyulas, C., Noy, N. F., & Musen, M. A. (2013). WebProtégé: A collaborative ontology editor and knowledge acquisition tool for the web. *Semantic Web*, *4*(1), 89–99. https://doi.org/10.3233/sw-2012-0057, PubMed: 23807872

Vergoulis, T., Chatzopoulos, S., Dalamagas, T., & Tryfonopoulos, C. (2020). Veto: Expert set expansion in academia. In M. Hall, T. Merčun, T. Risse, & F. Duchateau (Eds.), *Digital libraries for open knowledge* (pp. 48–61). Cham: Springer. https://doi.org/10.1007/978-3-030-54956-5_4

Vrandečić, D., & Krötzsch, M. (2009). Semantic MediaWiki. In J. Davies, M. Grobelnik, & D. Mladenić (Eds.), *Semantic knowledge management: Integrating ontology management, knowledge discovery, and human language technologies* (pp. 171–179). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-88845-1_13

Widdows, D., Peters, S., Cederberg, S., Chan, C.-K., Steffen, D., & Buitelaar, P. (2003). Unsupervised monolingual and bilingual word-sense disambiguation of medical documents using UMLS. In *Proceedings of the ACL 2003 Workshop on Natural Language Processing in Biomedicine* (pp. 9–16). Sapporo, Japan: Association for Computational Linguistics. https://doi.org/10.3115/1118958.1118960

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2024). *Experimentation in software engineering* (2nd ed.). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-69306-3

Yan, E. (2014). Research dynamics: Measuring the continuity and popularity of research topics. *Journal of Informetrics*, *8*(1), 98–110. https://doi.org/10.1016/j.joi.2013.10.010

Yang, H., & Lee, H. J. (2018). Research trend visualization by MeSH terms from PubMed. *International Journal of Environmental Research and Public Health*, *15*(6), 1113. https://doi.org/10.3390/ijerph15061113, PubMed: 29848974

Yip, H. Y., Nguyen, V., & Bodenreider, O. (2019). Construction of UMLS metathesaurus with knowledge-infused deep learning. In *BlockSW/CKG@ISWC*. https://ceur-ws.org/Vol-2599/CKG2019_paper_3.pdf

Zaharee, M. (2013). Building controlled vocabularies for metadata harmonization. *Bulletin of the American Society for Information Science and Technology*, *39*(2), 39–42. https://doi.org/10.1002/bult.2013.1720390211

Zapilko, B., Schaible, J., Mayr, P., & Mathiak, B. (2013). TheSoz: A SKOS representation of the thesaurus for the social sciences. *Semantic Web*, *4*(3), 257–263. https://doi.org/10.3233/SW-2012-0081

Zeng, M. L. (2008). Knowledge organization systems (KOS). *Knowledge Organization*, *35*(2–3), 160–182. https://doi.org/10.5771/0943-7444-2008-2-3-160

Zeng, M. L., & Mayr, P. (2019). Knowledge organization systems (KOS) in the semantic web: A multi-dimensional review. *International Journal on Digital Libraries*, *20*(3), 209–230. https://doi.org/10.1007/s00799-018-0241-2

Zhang, C., Tao, F., Chen, X., Shen, J., Jiang, M., … Han, J. (2018). TaxoGen: Unsupervised topic taxonomy construction by adaptive term embedding and clustering. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2701–2709). https://doi.org/10.1145/3219819.3220064

Zhang, X., Chandrasegaran, S., & Ma, K.-L. (2021). ConceptScope: Organizing and visualizing knowledge in documents based on domain ontology. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Article 19). https://doi.org/10.1145/3411764.3445396