A Survey on 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, United Kingdom**
[2]**CNR-ISTI — National Research Council, Institute of Information Science and Technologies "Alessandro Faedo", 56124 Pisa, Italy**
[3]**Department of Business and Law, University of Milano Bicocca, Milan, Italy**
[*]**Corresponding author: angelo.salatino@open.ac.uk**

ABSTRACT

Knowledge Organization Systems (KOSs), such as term lists, thesauri, taxonomies, and ontologies, play a fundamental role in categorising, 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 *i)* enhance retrievability of relevant documents, *ii)* enable advanced analytic solutions to quantify the impact of academic research, and *iii)* analyse and forecast research dynamics. This paper aims to present a comprehensive survey of the current KOS for academic disciplines. We analysed 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.

Keywords: Knowledge Organization Systems, Controlled Vocabularies, Taxonomies, Scholarly Ontologies, Scholarly Knowledge, Digital Libraries

1 INTRODUCTION

Knowledge Organization Systems (KOSs), such as term lists, thesauri, taxonomies, and ontologies, play a fundamental role in categorising, 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 and Lariviére, 2018), with the primary aim of classifying research articles, books, courses, patents, grants, software, experiment materials, scientific venues, domain experts, organisations, and several other relevant items and agents. These structured representation of knowledge have been adopted by most academic fields and proved very effective in *i)* improving the retrievability of relevant documents (Newman et al., 2010; Salatino et al., 2019), *ii)* enabling advanced analytic solutions to quantify the impact of research (Ding et al., 2014; Qiu et al., 2017; Sugimoto and Lariviére, 2018; Salatino et al., 2023), and *iii)* understanding and forecasting research dynamics (Scharnhorst et al., 2012; Qiu et al., 2017).

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 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 being "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 et al., 2020) and semi-automating systematic reviews (Bolanos et al., 2024). Despite the emergence of new AI systems based on Large Language Models (LLMs) in the last two 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 (Béchard and Ayala, 2024; Gnoli 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 categorise research products and are routinely adopted by libraries, online repositories, researchers, and organisations. Other fields (e.g., *"Mathematics"*) prevalently rely on one widely accepted KOSs (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 analysed 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 the **curation**, which includes information about the formats, the license, the frequency of the updates, the procedure used for the 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 characterisation 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 analyse how current solutions could be integrated and interlinked to produce a more comprehensive and granular representation of all academic disciplines.

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

The rest of this manuscript is organised 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 summarising the contributions and the main findings.

2 BACKGROUND

Knowledge organization systems play a crucial role in research by providing structured frameworks for organising complex information, allowing researchers to establish clear categories, discern relationships, and navigate large datasets 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 categorising and organising knowledge, retrieving information, facil-

[1] Full table describing the analysed KOSs according to the 15 features - https://doi.org/10.48366/R732033

[2] The code for processing the analysed KOSs is available on GitHub - https://github.com/angelosalatino/kos-rf

itating 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 emphasises 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 i) disambiguation of terms, ii) management of synonyms or equivalent terms, iii) establishment of semantic relationships, particularly hierarchical and associative links, and iv) 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 organising 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 system to provide full text of documents linked to bibliographic records. Major bibliographic databases like Web of Science, Scopus, Dimensions, and OpenAlex also employ KOSs to organise their vast collections of bibliographic records.

The annotation of documents based on the concepts within KOSs can be either performed 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. 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 et al., 2002), recommender systems (Cleverley and Burnett, 2015), integration and interoperability (Zeng and Mayr, 2019), and knowledge management and preservation (Chowdhury, 2010; Kopácsi et al., 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 et al., 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 personalised recommendations derived from user search patterns, fostering serendipitous discovery and richer user engagement (Cleverley and Burnett, 2015; Thanapalasingam 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 and Mayr, 2019).

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

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

2.2 Knowledge Organization Systems in Information Science

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

In the context of analysing 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, organisations, and countries. The dashboard leverages the Computer Science Ontology (Salatino et al., 2018b) 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 on potential research questions, and unexplored research avenues. Furthermore, the work of Yang and Lee (2018) introduced a tool that assists users in analysing research trends by expanding initial queries using MeSH terms, facilitating a more comprehensive exploration of the research landscape.

For research trends analysis, Ilgisonis et al. (2022) performed a systematic retrospective analysis of the frequencies of MeSH concepts across twelve years. Their analysis revealed potential shifts in scientific priorities, and they employed the same patterns to predict emerging trends within a five-year timeframe. In addition, Ovalle-Perandones 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 et al., 2018b) were employed to model the researcher's expertise. Specifically, they characterise the expertise of an author at the time of collaboration as the distribution of research topics of their paper over the preceding five years. Additionally, Kang et al. (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 and Ahlgren, 2020) to systematically categorise topics and disciplines within their dataset. Their topic-based analysis yields compelling insights that could shape future research policy decisions. Moreover, Chatzopoulos 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.

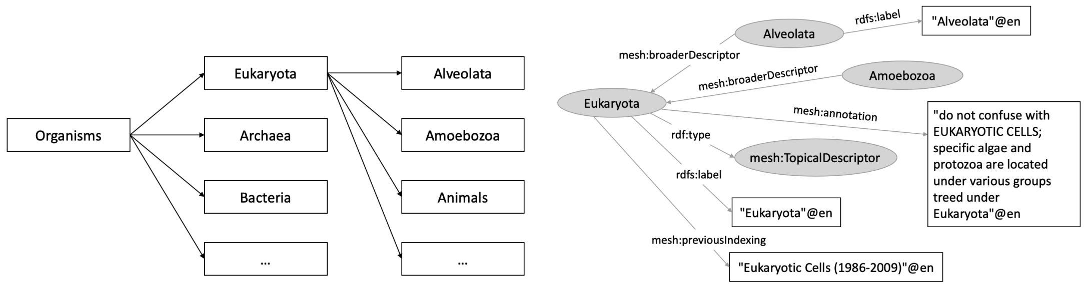
These analyses and approaches underscore the value and importance of KOSs in research, as they provide the essential structure and organisation needed for developing various downstream applications.

3 SURVEY METHODOLOGY

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

3.1 Concepts

In this survey, we focus on KOSs of academic fields. KOSs are commonly used to categorise items into specific classes. In this context, classes correspond to research topics within various disciplines. The items categorised can include a wide range of artefacts, such as documents (e.g., research papers, patents, project reports), videos (e.g., university courses), datasets, 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



(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.

"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 **research field** or **discipline** as a broad area of knowledge within academia, which 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 categories: *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 which support the organisation 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), whereas within the ISO 25964-2, it is referred to as "terminology" (BS ISO 25964-2:2013). 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*) organises the classes in a hierarchical structure with parent-child relationships (Rasch, 1987). In practical terms, as shown in Fig. 1a, a taxonomy is typically organised in a tree structure, with a root node at the top that unfolds in several sub-branches. An important characteristic is that all items classified according to a class can also be considered under all its super-classes. For instance, in Fig. 1a, all *"Bacteria"* are also *"Organisms"*.

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

Finally, an **ontology**, from the *"Information Technology"* perspective, is a formal and explicit description of knowledge within a domain, categorising things according to their essential or relevant qualities (Gruber, 1993). Ontologies conceptualise a subject area according to an abstract model of how domain experts rationalise knowledge in the domain. In practice, ontologies consist of a set of concepts, objects, and other entities and the relationships between them (Genesereth and Nilsson, 2012). Compared to thesauri, ontologies are more expressive since they can also express axioms and restrictions, so to provide local constraints on properties. As an example, Fig. 1b shows a portion of the *MeSH* ontology. In particular, the two `mesh:broaderDescriptor` relationships define the super-class 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 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

[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

disorganised. 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 and 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 Sec. 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 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.

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 organising 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)*".

Due to Google Scholar's limitations with logical operators, we executed the queries individually and then combined the resulting papers. In this phase, we gathered: i) *WikiCSSH* (Han et al., 2020), ii) *Computer Science Ontology* (Salatino et al., 2018b), iii) *Medical Subject Headings* (Darmoni et al., 2012), iv) *Library of Congress Classification* (Summers et al., 2008), v) *Modern Science Ontology* (Fathalla et al., 2023), vi) *Science Metrix Classification* (Archambault et al., 2011), and vii) *OpenAIRE's Field of Science Taxonomy* (Kotitsas et al., 2023).

Phase 2: We analysed the websites of academic publishers, pre-print archives, and academic search engines to identify the KOSs they use to organise their content. Specifically, we considered Scopus, Web of Science, Microsoft Academic Graph, Dimensions.ai, PubMed, the Springer Nature portal, ACM, IEEE, OpenAlex, and ArXiv. As a result, we identified: i) *ACM Computing Classification Scheme*, ANZSRC ii) *Fields of Research*, iii) *ArXiv Subjects*, Microsoft's, iv) *Fields of Study*, v) *Nature Subjects*, vi) *IEEE Thesaurus*, vii) *Web of Science Categories*, viii) *All Science Journal Classification Codes* of Scopus, and ix) *OpenAlex Topics*.

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

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: i) *Art and Architecture Thesaurus*, ii) *Biomedical Ontologies from BioPortal*, iii) *ChemOnt*, iv) *GeoRef Thesaurus*, v) *KNOWMAK*, vi) *Journal of Economic Literature*, vii) *Mathematics Subject Classification*, viii) *U.S. Geological Survey Library Classification System*, ix) *PhilPapers Taxonomy*, x) *Physical Subject Headings*, xi) *PsycInfo and PsycTests Classification Systems*, and xii) *Subject Resource Application Ontology*.

Phase 4: We contacted researchers working in the field of *"Digital Libraries"* and asked them whether they could point us toward any additional effort. As a result, we gathered: i) *Dewey Decimal Classification* and ii) *Unified Medical Language System*.

Phase 5: We expanded the resulting KOSs by analysing their links on Wikipedia. Specifically, we relied on the "See also" section, which typically lists online databases and other related KOSs. As a result, we gathered: i) *STW Thesaurus for Economics*, ii) *Physics and Astronomy Classification Scheme*, iii) *Open Biological and Biomedical Ontology*, iv) *National Library of Medicine classification*, and v) *Fields of research and development*.

Phase 6: We retrieved and considered all the KOSs that are explicitly linked into the already selected ones. Thanks to this process, we included: i) *Agrovoc Thesaurus*, ii) *DFG Classification*, iii) *EDAM*, iv) *TheSoz*, ANZSRC's v) *Socio-Economic Objective*, ASRC's vi) *Research Fields, Courses and Disciplines*, vii) *EuroVoc*, and viii) *Unesco Thesaurus*.

Phase 7: Whenever we could not find at least one KOS within one of the 19 disciplines defined in IC2 (e.g., *"History"*), we reached out to professors in the missing disciplines, asking which KOS they employ, if any. Their response directed us toward multi-field KOS such as i) *European Commission Taxonomy* and ii) *European Research Council Taxonomy*.

3.4 Methodology for the analysis of KOSs

We analysed each KOS according to the five main aspects summarised in Figure 2: *Scope*, *Structure*, *Curation*, *External Links*, and *Usage*.

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* that mainly focuses on *"Philosophy"*, but also describes some concepts from other fields, including *"Mathematics"*. Another category of KOSs covers multiple fields by design, typically because they aim to offer good coverage of the full set of scientific or academic disciplines. A good example is the *Unesco Thesaurus*, a well-known knowledge organization system which aims to cover all academic disciplines and is adopted by several libraries.

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 Sec. 3.1: term lists, taxonomies, thesauri, and ontologies. For example, the *Web of Science Categories* is a flat list of terms, and thus, we characterised it as a term list. The *Mathematics Subject Classification* and ANZSRC's *Fields of Research* are taxonomies since 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 may be 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, since deep taxonomies tend to include a granular representation of very detailed (narrow) topics and may allow for more fine-grained content organisation.

If the KOS is hierarchical, we also characterised it as either mono-hierarchical or poly-hierarchical according to the type of employed hierarchy. In mono-hierarchical 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 organise their concepts within a strict mono-hierarchical

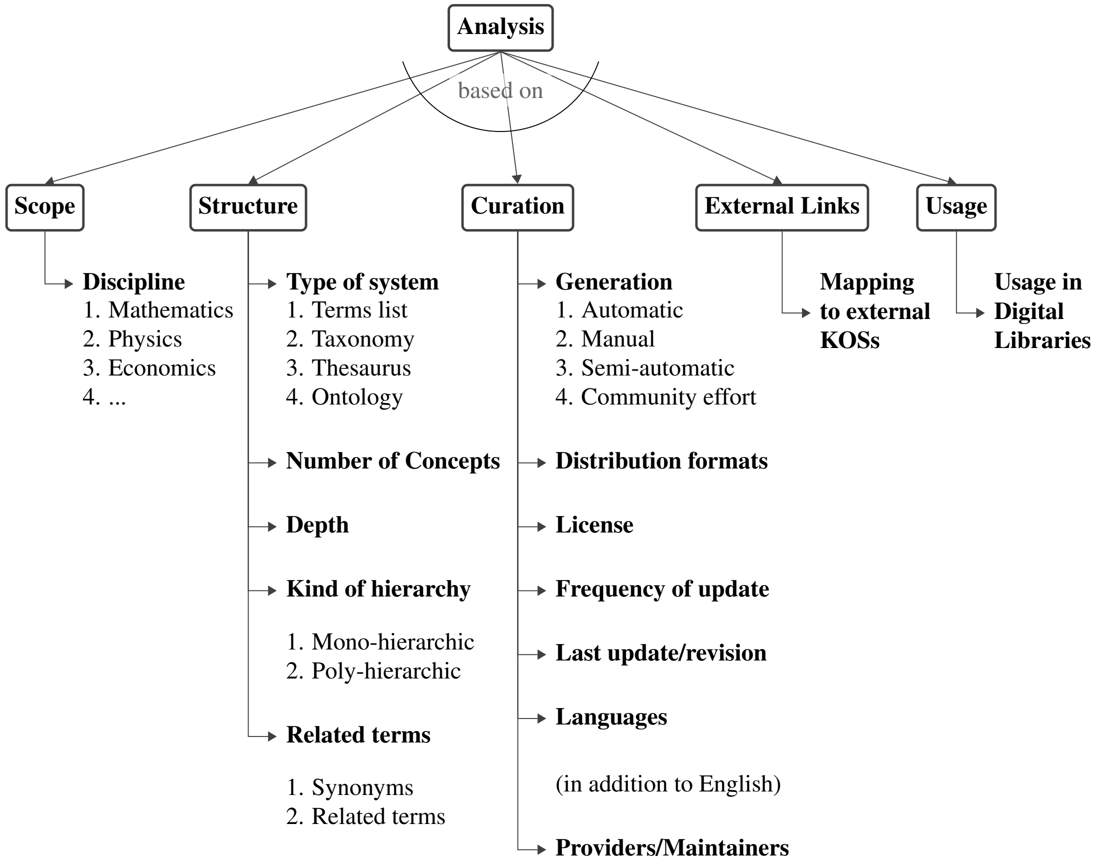


Figure 2. The features used for the analysis of knowledge organization systems.

structure (e.g., the *Mathematics Subject Classification*), while others (e.g., *STW Thesaurus for Economics*, *Medical Subject Headings*) are organised 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.

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, semi-automatic, 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 *semi-automatic* category covers KOSs generated semi-automatically. For instance, the *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

[10] Jupyter notebooks for analysing KOSs - https://github.com/angelosalatino/kos-rf

(TTL), 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-readable 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 these curatorial features have been identified from the official websites and relevant literature.

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, the *JEL classification*, and others. We will discuss the crucial implications of this mechanism in Sec. 5.

Usage. The final aspect that we considered regards the presence of collections of documents or other artefacts that are annotated or organised according to the KOS. Their existence attests to the fact that a portion of the community actively uses KOS. Furthermore, annotated datasets can also be used for training machine learning classifiers able to categorise new documents according to the original KOS (Kandimalla et al., 2021; Salatino et al., 2021). For instance, the *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 analyses the KOSs of research fields utilising the 15 features reported in Fig. 2. We identified 45 KOSs: 23 specialising in single disciplines and 22 covering multiple fields. Table 1 summarises 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 analysed features is available at https://doi.org/10.48366/R732033.

The following subsections discuss each of the five aspects outlined in Section 3.4, highlighting trends and providing specific examples.

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 KOS, 22 cover multiple fields, while 23 focus on a single field of science. However, only 12 of the 19 fields under analysis are addressed by at least one single-field KOS, while the remaining ones rely exclusively on broader multi-field solutions.

Notably, 6 fields out of the 19 under analysis (see IC 2, Sec. 3.2) are covered by more than one KOS. In *"Medicine"*, we identified four single-field KOSs: the *Medical Subject Headings*, the *Unified Medical Language System*, the *National Library of Medicine classification*, and the *Biomedical Ontologies from*

[11] Resource Description Framework (RDF) — https://www.w3.org/RDF
[12] OBO Format — http://purl.obolibrary.org/obo/oboformat/spec.html
[13] Creative Commons — https://creativecommons.org
[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 on the 45 knowledge organization systems, reporting their main discipline, number of concepts, depth, type of system (**Sy** = **O**ntology, **T**axonomy, Thesa**U**rus, and Term **L**ist), kind of hierarchy (**Hr** = **P**oly-hierarchical or **M**ono-hierarchical), and related terms (**RT** = **Y**es or **N**o).

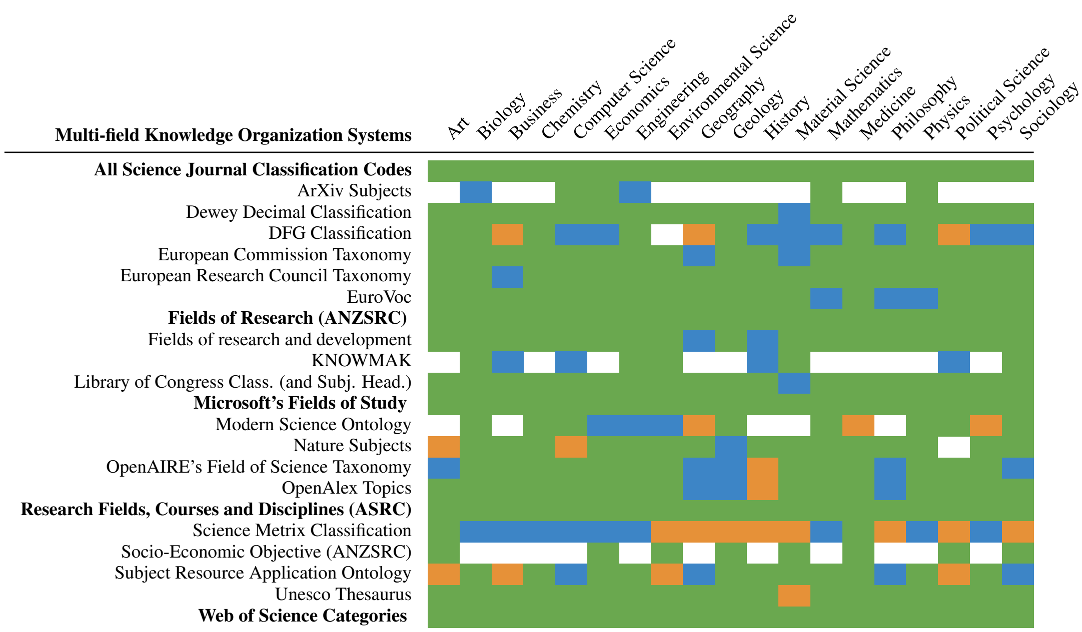
Knowledge Organization System	Main Discipline	#Concepts	Depth	Sy	Hr	RT
Agrovoc Thesaurus	Agriculture	41K	14	O	P	Y
Art and Architecture Thesaurus	Art & Architecture	58K	13	O	P	Y
EDAM	Bioinformatics	264	7	O	P	Y
Open Biological and Biomedical Ontology	Biology	5.3M	39	O	P	Y
Biomedical Ontologies from BioPortal	Biomedicine	13M	-	O	P	Y
ChemOnt	Chemistry	4825	11	O	M	Y
ACM Computing Classification Scheme	Computer Science	2114	6	T	P	N
Computer Science Ontology	Computer Science	14K	13	O	P	Y
Computer Science Subject Headings from Wikipedia	Computer Science	7354	20	T	M	N
Journal of Economic Literature	Economics	859	3	T	M	N
STW Thesaurus for Economics	Economics	14K	13	O	P	Y
IEEE Thesaurus	Engineering	12K	3	U	P	Y
GeoRef Thesaurus	Geology	33K	5	U	M	Y
U.S. Geological Survey Library Classification System	Geology	2k	6	T	M	N
Mathematics Subject Classification	Mathematics	6598	3	T	M	Y
Medical Subject Headings	Medicine	30K	13	O	P	Y
National Library of Medicine classification	Medicine	4761	4	T	M	N
Unified Medical Language System	Medicine	3.3M	30	U	P	Y
PhilPapers Taxonomy	Philosophy	7447	5	T	P	N
Physical Subject Headings	Physics	3749	6	O	P	Y
Physics and Astronomy Classif. Scheme	Physics	8260	5	T	M	N
PsycInfo and PsycTests Classification Systems	Psychology	188	3	T	M	N
TheSoz	Social Science	8223	6	O	P	Y
All Science Journal Classification Codes	Multi-field	334	2	T	M	N
ArXiv Subjects	Multi-field	176	3	T	M	N
Dewey Decimal Classification	Multi-field	60K	14	T	P	Y
DFG Classification	Multi-field	278	4	T	M	N
European Commission Taxonomy	Multi-field	629	3	T	M	N
European Research Council Taxonomy	Multi-field	431	3	T	M	N
EuroVoc	Multi-field	7439	6	O	P	Y
Fields of Research (ANZSRC)	Multi-field	4406	3	T	M	N
Fields of research and development (OECD)	Multi-field	48	2	T	M	N
KNOWMAK	Multi-field	72	3	O	P	Y
Library of Congress Class. (and Subj. Head.)	Multi-field	467K	29	O	P	Y
Microsoft's Fields of Study	Multi-field	704K	6	T	P	N
Modern Science Ontology	Multi-field	369	6	O	P	N
Nature Subjects	Multi-field	2852	8	O	P	N
OpenAIRE's Field of Science Taxonomy	Multi-field	50K	6	T	P	N
OpenAlex Topics	Multi-field	4798	4	T	M	N
Research Fields, Courses and Disciplines (ASRC)	Multi-field	1061	3	T	M	N
Science Metrix Classification	Multi-field	199	3	T	M	N
Socio-Economic Objective (ANZSRC)	Multi-field	1974	3	T	M	N
Subject Resource Application Ontology	Multi-field	435	8	O	P	Y
Unesco Thesaurus	Multi-field	4482	6	O	P	Y
Web of Science Categories	Multi-field	254	1	L	-	N

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 also rely on the *Physics and Astronomy Classification Scheme* and *Physical Subject Headings*.

Most other fields rely on just one KOS. This is the case for *"Agriculture"* (*Agrovoc Thesaurus*), *"Art and Architecture"* (*Art and Architecture Thesaurus*), *"Chemistry"* (*ChemOnt*), *"Engineering"* (*IEEE Thesaurus*), *"Mathematics"* (*Mathematics Subject Classification*), *"Philosophy"* (*PhilPapers Taxonomy*), and *"Psychology"* (*PsycInfo and PsycTests Classification Systems*).

Four KOSs were not directly assigned to one of the standard 19 broad fields, as they focus on more specialised areas. Specifically, we associated *Agrovoc* to *"Agriculture"*, which is a sub-area 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 super-area of *"Economics"*,

Table 2. Coverage of the multi-field 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.



"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 multi-field 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 multi-field KOSs, we observed a substantial diversity in topic coverage. Table 2 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 sub-areas. Notably, only five of these KOSs, emphasised in bold within the table, consistently cover all the 19 top-level research areas presented in Section 3.

4.2 Structure

This section analyses the KOSs based on the previously defined structural features, focusing on their size, depth, type, hierarchical organisation, 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 categorising 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, 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 and 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, seventeen 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 analysed KOSs is approximately 4,700. Sixteen single-field KOSs exceed this median in concept count, while sixteen of the multi-field KOSs contain fewer concepts. This pattern suggests a trend: single-field KOSs tend to be larger and more specialised, likely to capture the intricacies within their specific domains. In contrast, multi-field KOSs seem to be designed to offer a broader overview across various fields, often resulting in a smaller number of concepts.

The six multi-field 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* (50K), *Dewey Decimal Classification* (60K), *Library of Congress Classification* (467K), and *Microsoft's Fields of Study* (704K).

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*). In particular, 15 KOSs employ up to 3 levels, 15 use from 4 to 6 levels, 3 utilise from 7 to 10 levels, and 11 extend beyond 10 levels.

The median depth stands at 6, with 18 of the 22 multi-field KOSs falling below this threshold. This finding further supports the notion that the majority of multi-field systems aim to provide a broad overview across diverse research fields.

4.2.4 Kind of hierarchy

Column 'Hr' in Table 1 reveals a near-equal distribution between poly-hierarchical (24) and mono-hierarchical (20) KOSs. Only one KOS (*Web of Science Categories*) is considered non-hierarchical since it is a flat list of terms. When examining the 23 single-field KOSs, it is interesting to note that poly-hierarchical structures (14) outnumber mono-hierarchical ones (9). In contrast, multi-field systems exhibit a nearly equal distribution, with 11 mono-hierarchical and 10 poly-hierarchical structures. This trend suggests that single-field KOSs are more likely to employ poly-hierarchical structures, as they need to capture fine-grained research topics that may be derived 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 and 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). 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 categorised as manual, automated, semi-automated, or community-based.

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

Six KOSs have already adopted semi-automatic methodologies: *Biomedical Ontologies from BioPortal*, *KNOWMAK*, *Microsoft's Fields of Study*, *OpenAIRE's Field of Science Taxonomy*, *Subject Resource*

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

Application Ontology, and *OpenAlex Topics*. For instance, *OpenAlex Topics* combines manually curated concepts from the *All Science Journal Classification Codes* 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 labelled 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 and Motta, 2015), which processed 16 million scientific publications. Klink-2 identifies relationships between topics by analysing 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 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] (i.e., 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 lacks the structured organisation necessary for direct and automated processing by computer systems. Several KOSs are available in machine-readable formats: 17 utilise the RDF standard, and 10 are available in CSV format. Finally, 6 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.

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

[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 in the upper part and multi-field KOSs 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
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		

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 analysing 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 the *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*, *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*, *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, the *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. On the contrary, the *Fields of Research (ANZSRC)* consists of only three levels, and since the modelled 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 the most recent update in 2020, after 12 years.

We also investigated whether the frequency of updates is related to the type of system. Our analysis revealed that among the 18 KOSs updated within a year, the majority were ontologies (11), followed by taxonomies (5) and thesauri (2). In contrast, the 8 KOSs that were either updated every ten years or discontinued were taxonomies. This suggests that while taxonomies might be effective initially for organising subjects hierarchically, they may become increasingly challenging 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

Seventeen systems have received their most recent updates in 2023 or later, while thirteen were last updated between 2020 and 2022. This indicates that within the last five years, thirty KOSs have been revised to reflect the evolving nature of their respective fields. Of these, 14 are single-field systems, and 15 are multi-field.

The remaining fifteen 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 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 had at least one version in English. Thirty-five KOSs are available exclusively in English. Conversely, ten KOSs provide research topics in multiple languages, driven by their intended use and jurisdictional context. For example, *EuroVoc*, developed by the Publications Office of the European Union, supports 23 languages to facilitate seamless interoperability within European digital libraries. The additional nine multilingual KOSs are *Agrovoc Thesaurus*, *Dewey Decimal Classification*, *Fields of research and development*, *STW Thesaurus for Economics*, *Unified Medical Language System*, *TheSoz*, *Science Metrix Classification*, *Unesco Thesaurus* and *Art and Architecture Thesaurus*.

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* (5 languages), and the *Fields of research and development* (6 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 Fig. 3, *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. 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.

Finally, an analysis of the 10 multilingual KOSs reveals that the majority (6) are ontologies, while 3 are taxonomies, and 1 is a thesaurus. This trend likely reflects the inherent capacity of ontologies to effectively manage concept labels across multiple languages (Montiel-Ponsoda et al., 2011).

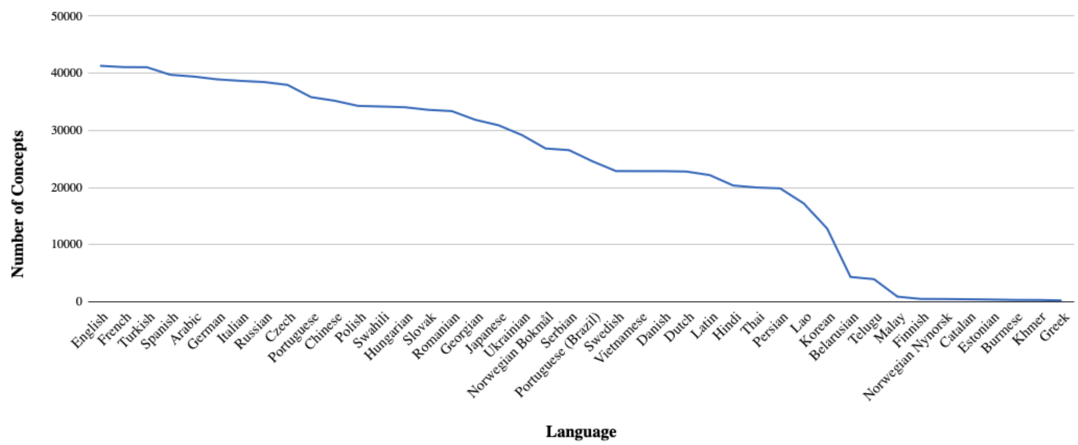


Figure 3. Distribution of terms per language in the Agrovoc Thesaurus. Produced with term counts available on https://agrovoc.fao.org/browse/agrovoc (Agrovoc v.2024-04, Apr. 2024).

4.3.7 Providers and Maintainers

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

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 organise, 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 organisations 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 (UK), University of Sheffield (UK), University of Illinois at Urbana-Champaign (USA), University of Alberta (CA), Athena Research Center (GR), and The Getty Research Institute (USA) produced 8 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 7 KOS.

Finally, only 2 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 multi-field KOSs (9 out of 12), likely due to their interdisciplinary focus. In contrast, Research Institutes & Universities predominantly created single-field KOSs (6 out of 8), possibly because research teams tend to specialise in a specific area. The remaining organisational categories demonstrated a balanced approach in developing both single and multi-field KOSs.

4.4 External Links

Eighteen 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-know 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, such as Wikidata or DBpedia, 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.

Wikidata[23] stands out as the most externally linked knowledge graph since 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*, *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.

Twelve out of the initial eighteen earn a five-star rating within the Linked Open Data deployment scheme proposed by Sir Tim Berners-Lee in 2010[24]. Their extensive interconnections, facilitated by RDF standards such as OWL and SKOS, establish a rich network of information with related KOSs. Adhering to the 5-star scheme[25] optimises 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 organising digital libraries, enhancing metadata, and ensuring interoperability between different data systems. For example, the *Fields of Research (ANZSRC)* is utilised by Dimensions.ai to organise research metadata, by Figshare to manage research repositories, and by the Australian and New Zealand governments to measure and analyse research and experimental development.

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

Twenty-nine KOSs are currently being employed to organise digital libraries including ACM Digital Library, IEEE Digital Library, MEDLINE/PubMed, Scopus, Nature, EconLit, Physical Review, and Mathematical Reviews. Ten KOSs are employed to organise 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 organise 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 KOS are also utilised across a wide array of initiatives to directly advance research and create additional knowledge and tools. For instance, the *Computer Science Ontology* has been employed

[21] Sunflowers (plural) in Agrovoc Thesarus - http://aims.fao.org/aos/agrovoc/c_aad037e4
[22] Sunflower (singular) in Eurovoc - http://eurovoc.europa.eu/4472
[23] Wikidata - https://www.wikidata.org
[24] Linked Data - https://www.w3.org/DesignIssues/LinkedData.html
[25] Linked Open Data 5 Star - https://www.w3.org/DesignIssues/LinkedData.html

for exploring and analysing scholarly data through platforms like Rexplore (Osborne et al., 2013), ScholarLensViz (Löffler et al., 2020), ConceptScope (Zhang et al., 2021), and VeTo (Vergoulis et al., 2020). It has also been instrumental in generating knowledge graphs, such as AIDA KG (Angioni et al., 2022b) and CS KG (Dessì et al., 2022), as well as in recommending video lessons (Borges and dos Reis, 2019). However, thoroughly analysing 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 Sec. 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. Sec. 5.2 discusses the most important limitations and challenges regarding the integration of KOSs. We then present several other important challenges characterising this space, including: *i)* expanding the coverage of different languages (Sec. 5.3), *ii)* reconciling expert disagreements (Sec. 5.4), *iii)* assessing the quality of KOSs (Sec. 5.5), *iv)* handling ambiguous labels (Sec. 5.6), *v)* improving the generation and frequency of updates (Sec. 5.7), and *vi)* developing scalable and accurate approaches for item classification (Sec. 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 multi-field 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 simultaneously is: i) **comprehensive**, i.e., it encompasses all the main academic fields, ii) **granular**, i.e., it represents all the specific research areas that are typically used by researchers to refer to their work, iii) **maintained**, i.e., it is constantly updated to reflect the latest developments, and iv) **open**, i.e., it can be used freely, without restrictions.

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

A potential approach to creating this resource is to adopt one or more multi-field 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.

Multi-field KOSs that may support the creation of a general KOS. It is useful to analyse to what extent the current multi-field KOSs fulfil the requirements defined in the previous session. Table 4 presents an evaluation of the 22 multi-field KOSs examined in this study, considering four key features: comprehensiveness, depth, frequency of updates, and openness. For comprehensiveness, colours are assigned based on the KOS coverage, detailed by Table 2. Specifically, green indicates full coverage, orange denotes partial coverage, and red signifies that the KOSs cover only a limited set of fields. In the depth column, the colour coding is as follows: red indicates a depth of 2 or less (low or non-existent), 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., every 5–10 years), and green denotes more frequent updates (i.e., within 2 years). If such information is unavailable, it is marked in white. Lastly, in the openness column, red indicates that the KOS is copyrighted and requires a license, orange signifies that the KOS can be used with certain restrictions, and green shows that it is an open resource.

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)*, *Micorsoft's Fields of Study*, and *Web of Science Categories*. However, none of these KOSs fulfils 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 the *Microsoft's Fields of Study* offer more depth compared to the previous two, but they were discontinued. The *Fields of*

Table 4. Analysis of KOSs according to their comprehensiveness, depth, frequency of update, and openness. Features of each KOSs are evaluated using a colour-coded system: ■ green indicates a strong performance, ■ orange signifies an adequate performance, and ■ red reflects poor performance. A feature is blank when the information is not available.

	Comprehensiveness	Depth	Frequency of update	Openness
All Science Journal Classification Codes	green	red		red
ArXiv Subjects	red	orange		green
Dewey Decimal Classification	orange	green	green	red
DFG Classification	red	orange		green
European Commission Taxonomy	orange	orange		green
European Research Council Taxonomy	orange	orange		green
EuroVoc	orange	green	green	green
Fields of Research (ANZSRC)	green	orange	orange	green
Fields of research and development	orange	red		red
KNOWMAK	red	orange		orange
Library of Congress Class. (and Subj. Head.)	orange	green	green	red
Microsoft's Fields of Study	green	green	red	green
Modern Science Ontology	red	green		green
Nature Subjects	orange	green		red
OpenAIRE's Field of Science Taxonomy	orange	green	green	green
OpenAlex Topics	orange	orange	green	green
Research Fields, Courses and Disciplines (ASRC)	green	orange	red	green
Science Metrix Classification	orange	orange	orange	green
Socio-Economic Objective (ANZSRC)	red	orange	orange	green
Subject Resource Application Ontology	orange	green	green	green
Unesco Thesaurus	orange	green	green	orange
Web of Science Categories	green	red		red

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 for using them.

In conclusion, the current state-of-the-art multi-field 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 multi-field KOSs, incorporating missing concepts from specialised single-field KOSs like *Mathematics Subject Classification*, *PhilPapers Taxonomy*, and *Physical Subject Headings*.

Academic fields that are not currently covered by specific KOSs. Several disciplines lack dedicated KOS and are only superficially addressed by multi-field systems, making it challenging to categorise documents with the necessary precision. In particular, we were unable to identify any KOS that offers a good characterisation of seven research fields: *"History"*, *"Political Science"*, *"Environmental Science"*, *"Material Science"*, *"Geography"*, *"Sociology"*, and *"Business"*. In the fields of *"Political Science"* and *"Sociology"*, *The Soz* could potentially be utilised. However, *The Soz* 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 the creation of 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*, *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 categorised into the following sub-challenges:

Sub-challenge 1: Generating mappings between KOSs;

Sub-challenge 2: Adopting standard formats;

Sub-challenge 3: Developing tools for facilitating the integration of KOSs.

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 semi-automatic (Kalfoglou and 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 (Halper et al., 2011; Erdogan et al., 2010; Solimando et al., 2014).

Automatic approaches typically use a combination of similarity metrics, natural language processing, and machine learning (Salatino et al., 2020; Zapilko et al., 2013; Declerck, 2013). For instance, Salatino et al. (2020) associated research topics to the corresponding DBpedia entity with the DBpedia Spotlight API (Daiber et al., 2013). They fed the tool with artificial sentences listing the labels of the topic and of its direct sub- and super-topics, 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 have 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 and Euzenat, 2008; Solimando et al., 2014; Slater et al., 2020). 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 et al., 2011). Some recent approaches address these limitations by relying on description logic (Dhombres and Bodenreider, 2016) or deep learning (Yip et al., 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 for comparing 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 non-synonyms) and false negatives (fails to identify synonyms). The community still needs to further investigate and develop these new solutions in practical settings.

Finally, in *Semi-automatic approaches*, domain experts analyse, correct, and give feedback on candidate mappings produced by automatic approaches (Salvadores et al., 2013; Yip et al., 2019). It is worth mentioning two big endeavours 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[26]. 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.

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

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

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[27] (i.e., Resource Description Framework), which is a World Wide Web Consortium[28] (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[29], 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 et al., 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: <*http://www.nature.com/subjects/social-sciences*, *http://www.w3.org/2004/02/skos/core#narrower*, *http://www.nature.com/subjects/politics*>. Such representation allows AI systems to interconnect, identify, disambiguate, and integrate data effectively.

On the contrary, KOSs published in CSV or HTML are harder to treat since 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 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[30] KOSs. Examples include the RDFification of the *Mathematics Subject Classification*[31] and *IEEE Thesaurus*[32]. However, these RDF versions are not maintained by the original curators of the KOSs, making them difficult to update and reuse.

Developing tools for facilitating the integration of KOSs. The first KOSs were of relatively small size, such as the *Library Classification for Environmental Science*[33], 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[34] (DZHW) used Trello to build the *Research Core Dataset*[35] (RCD), a classification of interdisciplinary research fields (Stiller et al., 2021). Trello[36] 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,

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

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

[29] 5-star Open Data — https://www.w3.org/DesignIssues/LinkedData.html

[30] *RDFication* refers to the process of converting data into Resource Description Framework (RDF) format.

[31] MSC2020_SKOS (TIBHannover) - https://github.com/TIBHannover/MSC2020_SKOS

[32] ieee-taxonomy-thesaurus-rdf (The Open University) - https://github.com/angelosalatino/ieee-taxonomy-thesaurus-rdf

[33] 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.

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

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

[36] Trello — https://trello.com

building complex KOSs using Trello presents some limitations, since it does not allow users to nest cards in 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[37], Semantic MediaWiki[38] and PoolParty[39]. Protégé is an open-source software developed by the 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 et al., 2013).

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

PoolParty is a suite that supports the creation and maintenance of taxonomies, ontologies, knowledge graphs, and semantic search applications (Schandl and Blumauer, 2010). Similarly to WebProtégé, it allows users to browse and work collaboratively. Other commercial solutions include Data Harmony[40] and Synaptica[41] providing similar functionalities.

In brief, there is a wide collection of tools that can facilitate the integration of the different KOSs. However, all these tools are general-purpose and built to support a multitude of use cases. As a result, users that do not have experience with these technologies and semantic web technologies 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 and 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 and 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 only in five languages (English, Arabic, Russian, French, and Spanish). The *Agrovoc Thesaurus* has a very wide variety of languages; however, not all the forty-two languages are equally represented, as reported in Fig. 3. Notably, the Food and Agriculture Organization, who 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 analysed in this survey. We need further work to produce resources able to support multiple languages (Tudhope and 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 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 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

[37] Protégé — https://protege.stanford.edu
[38] Semantic MediaWiki — https://www.semantic-mediawiki.org
[39] PoolParty — https://www.poolparty.biz
[40] Data Harmony — https://www.accessinn.com/data-harmony
[41] Synaptica — https://www.synaptica.com

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 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 et al. (2019) analysed the agreement of experts in characterising 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 categorise UMLS concepts into nine different types. While the inter-annotator 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 behaviour and as a biological function. The authors emphasise 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* (Morrey et al., 2009; Ayele et al., 2012; Raad and 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. Fig. 4 shows the two examples, and specifically the arrows identify the `skos:narrower`[42] 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*[43]: *"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 confusion to the hierarchical structure which necessitates 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: *i)* ensuring correct hierarchical relationships between two subjects, and *ii)* addressing ambiguities in preserving multiple semantic views. Additional challenges are instead related to the *completeness* of the concepts (Lambe, 2014), such as: *i)* whether an area needs to be refined or enriched with sub-areas, *ii)* whether the description of concepts, including definition, scope, and editorial information, is appropriate, and *iii)* 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: *i)* gold

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

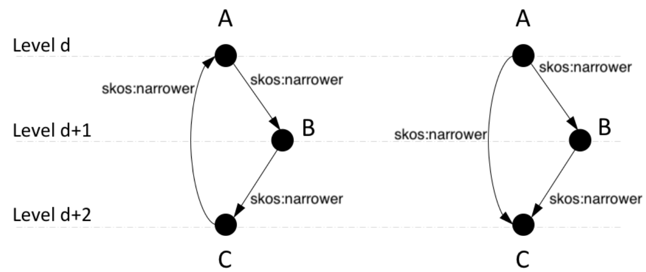


Figure 4. Example of inconsistencies within the KOSs. On the left the cyclic inconsistency and on the right the redundant inconsistency.

standard-based, *ii)* corpus-based, *iii)* task-based, and *iv)* 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 and 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 (Johnson et al., 2007; Djeddi and Khadir, 2014).

In the literature, we can find several approaches attempting to disambiguate the various senses of a term. Bella et al. (2017) mapped *Eurovoc* with *Universal Decimal Classification* (UDC)[44], 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* (Widdows et al., 2003; McInnes et al., 2007); however, these methods necessitate 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 et al. (2004) suggest to formally represent 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 et al. (2004) audited *UMLS* and found out 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, since the resulting KOSs are more complex (Shi et al., 2011), harder to maintain, and the type assignment may cause disagreements among domain experts (Fan and Friedman, 2007).

5.7 Automatic generation and updating KOSs

The current landscape of KOSs reveals gaps in coverage for certain areas (see Sec. 4.1) and highlights the vulnerability of some KOSs to becoming outdated (see Sec. 4.3.4). Given the critical role KOSs play in

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

various downstream services, the scientific community should prioritise 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 and Motta, 2015), or to scan the recent literature and extend existing ones, e.g., (Huang et al., 2020; Pisu et al., 2024). Klink-2 (Osborne and Motta, 2015) demonstrated high effectiveness in automatically generating large-scale and granular ontologies of research areas. Other approaches include Shang et al. (2020), who 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 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 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 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 (Revenko et al., 2024; Aggarwal 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 categorise a vast amount of items, such as articles, books, courses, patents, software, experiment materials, and so on. 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 categorises these approaches into two main types: supervised and unsupervised methods.

Supervised methods need to be trained according to a set of labelled 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 et al. (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 et al. (2015) trained their classifiers on a corpus of 3,186 papers distributed over 6 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 9M 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 labelled 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., 2021) 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 post-processes 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, however it tends to confuse polysemic terms. MetaMap (Aronson and 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 et al. (2010) developed cTAKES, which is an approach for extracting *UMLS* concepts from medical records. cTAKES implements a named-entity recognition approach, identifying *UMLS* terms within a noun-phrase look-up 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 generalise across diverse research areas (Kojima 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 et al. (2012). In the following discussion, we evaluate these potential threats and detail the measures we have taken to minimise their impact.

Internal Validity. Internal validity in surveys pertains to the rigour 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 organisation systems. The initial protocol was developed by the first author and subsequently reviewed and refined by co-authors to establish consensus before initiating the survey process. We utilised 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 multi-stage selection process to ensure a rigorous evaluation and minimise selection bias. Initially, the first author analysed 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 minimise 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.

External Validity. External validity refers to how well the results of this survey can be generalised and applied to other contexts and areas of study. To maximise the applicability of our findings, we utilised 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 the English version. This was set due to the predominance of English as the language of scientific communication (Sugimoto and Lariviére, 2018). Additionally, the absence of English versions would have hindered our ability to conduct in-depth analyses, such as the one presented in Table 2. 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 customised to fit the unique content of their respective libraries, which could result in a skewed portrayal of the broader scientific landscape.

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 analysed 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 fifteen specific features.

We recognise 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 specialised expertise across various scientific disciplines.

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 minimising 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 analysed a sample of KOSs using this standardised 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 categorisation, 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 overview of 45 KOSs, with 23 focusing on a single field and 22 covering multiple fields. We propose an analysis framework that characterises 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 on 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 re-run 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, Materials 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 analysed in this study.

The analysed 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 analysed (24) are still traditional taxonomies. However, a growing number of more recent KOS (18) are now based on ontologies. Regarding hierarchical structures, there is a fairly balanced distribution: 24 KOSs utilise a poly-hierarchical structure, while 20 employ a mono-hierarchical one. Traditionally, Knowledge Organization Systems (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 semi-automatic 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 multi-field KOS that simultaneously is comprehensive in topic coverage, granular, consistently updated, and openly accessible (see Table 4). The creation of such a KOS could revolutionise the field by dramatically improving content organisation 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 policy making.

We have identified key future research directions in this domain, along with the challenges that accompany them. These priorities include generating higher-quality multi-field KOSs, developing new automatic methods for integrating and updating KOSs, adopting standardised 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 capitalise 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 in this direction.

REFERENCES

Tanay Aggarwal, Angelo Salatino, Francesco Osborne, and Enrico Motta. Classifying scientific topic relationships with scibert. In *Sci-k 2024 @ ISWC 2024*. CEUR (in press), 2024.

Simone Angioni, Angelo Salatino, Francesco Osborne, Aliaksandr Birukou, Diego Reforgiato Recupero, and Enrico Motta. Leveraging knowledge graph technologies to assess journals and conferences at springer nature. In *International Semantic Web Conference*, pages 735–752. Springer, 2022a.

Simone Angioni, Angelo Salatino, Francesco Osborne, Diego Reforgiato Recupero, and Enrico Motta. AIDA: A knowledge graph about research dynamics in academia and industry. *Quantitative Science Studies*, 2(4):1356–1398, 02 2022b. ISSN 2641-3337. doi: 10.1162/qss_a_00162. URL https://doi.org/10.1162/qss_a_00162.

ANSI/NISO Z39.19-2005(R2010). Guidelines for the construction, format, and management of monolingual controlled vocabularies. Standard, National Information Standards Organization, Baltimore, Maryland, July 2010.

Éric Archambault, Olivier H Beauchesne, and Julie Caruso. Towards a multilingual, comprehensive and open scientific journal ontology. In *Proceedings of the 13th international conference of the international society for scientometrics and informetrics*, volume 13, pages 66–77. Durban, South Africa, 2011.

Alan R Aronson and François-Michel Lang. An overview of metamap: historical perspective and recent advances. *Journal of the American Medical Informatics Association*, 17(3):229–236, 2010.

Sören Auer, Viktor Kovtun, Manuel Prinz, Anna Kasprzik, Markus Stocker, and Maria Esther Vidal. Towards a knowledge graph for science. In *Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*, pages 1–6, 2018.

Demeke Ayele, Jean-Pierre Chevallet, Getnet Kassie, and Million Meshesha. Enhancing semantic relation quality of umls knowledge sources. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems*, MEDES '12, page 59–66, New York, NY, USA, 2012. Association for Computing Machinery. ISBN 9781450317559. doi: 10.1145/2457276.2457289. URL https://doi.org/10.1145/2457276.2457289.

Gábor Bella, Fausto Giunchiglia, and Fiona McNeill. Language and domain aware lightweight ontology matching. *Journal of Web Semantics*, 43:1–17, 2017. ISSN 1570-8268. doi: https://doi.org/10.1016/j.websem.2017.03.003. URL https://www.sciencedirect.com/science/article/pii/S1570826817300161.

Iz Beltagy, Kyle Lo, and Arman Cohan. Scibert: A pretrained language model for scientific text, 2019. URL https://arxiv.org/abs/1903.10676.

Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. *Scientific american*, 284(5):34–43, 2001.

Olivier Bodenreider. Circular hierarchical relationships in the umls: etiology, diagnosis, treatment, complications and prevention. In *Proceedings of the AMIA Symposium*, page 57. American Medical Informatics Association, 2001.

Olivier Bodenreider. The unified medical language system (umls): integrating biomedical terminology. *Nucleic Acids Research*, 32(suppl_1):D267–D270, 01 2004. ISSN 0305-1048. doi: 10.1093/nar/gkh061.

Francisco Bolanos, Angelo Salatino, Francesco Osborne, and Enrico Motta. Artificial intelligence for literature reviews: Opportunities and challenges, 2024. URL https://arxiv.org/abs/2402.08565.

Marcos Vinícius Macêdo Borges and Julio Cesar dos Reis. Semantic-enhanced recommendation of video lectures. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, volume 2161, pages 42–46. IEEE, 2019. doi: 10.1109/ICALT.2019.00013.

Jeffrey Brainard. Scientists are drowning in covid-19 papers. can new tools keep them afloat. *Science*, 13 (10.1126), 2020. doi: 10.1126/science.abc7839.

BS ISO 25964-2:2013. Information and documentation. thesauri and interoperability with othervocabularies. interoperability with other vocabularies. Standard, International Organization for Standardization, March 2013.

Patrice Béchard and Orlando Marquez Ayala. Reducing hallucination in structured outputs via retrieval-augmented generation, 2024. URL https://arxiv.org/abs/2404.08189.

Cornelia Caragea, Florin Bulgarov, and Rada Mihalcea. Co-training for topic classification of scholarly data. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 2357–2366, 2015.

Serafeim Chatzopoulos, Thanasis Vergoulis, Ilias Kanellos, Theodore Dalamagas, and Christos Tryfonopoulos. Artsim: Improved estimation of current impact for recent articles. In Ladjel Bellatreche, Mária Bieliková, Omar Boussaïd, Barbara Catania, Jérôme Darmont, Elena Demidova, Fabien Duchateau, Mark Hall, Tanja Merčun, Boris Novikov, Christos Papatheodorou, Thomas Risse, Oscar Romero, Lucile Sautot, Guilaine Talens, Robert Wrembel, and Maja Žumer, editors, *ADBIS, TPDL and EDA 2020 Common Workshops and Doctoral Consortium*, pages 323–334, Cham, 2020. Springer International Publishing. ISBN 978-3-030-55814-7.

Ekaterina Chernyak. An approach to the problem of annotation of research publications. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, WSDM '15, page 429–434, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450333177. doi: 10.1145/2684822.2697032. URL https://doi.org/10.1145/2684822.2697032.

Lydia B Chilton, Greg Little, Darren Edge, Daniel S Weld, and James A Landay. Cascade: Crowdsourcing taxonomy creation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1999–2008, 2013.

Gobinda Chowdhury. From digital libraries to digital preservation research: the importance of users and context. *Journal of documentation*, 66(2):207–223, 2010.

Paul H Cleverley and Simon Burnett. 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), 2015.

Suyang Dai, Ronghui You, Zhiyong Lu, Xiaodi Huang, Hiroshi Mamitsuka, and Shanfeng Zhu. Fullmesh: improving large-scale mesh indexing with full text. *Bioinformatics*, 36(5):1533–1541, 2020.

Joachim Daiber, Max Jakob, Chris Hokamp, and Pablo N. Mendes. Improving efficiency and accuracy in multilingual entity extraction. In *Proceedings of the 9th International Conference on Semantic Systems*, I-SEMANTICS '13, page 121–124, New York, NY, USA, 2013. Association for Computing Machinery. ISBN 9781450319720. doi: 10.1145/2506182.2506198. URL https://doi.org/10.1145/2506182.2506198.

Stéfan J Darmoni, Lina F Soualmia, Catherine Letord, Marie-Christine Jaulent, Nicolas Griffon, Benoît Thirion, and Aurélie Névéol. Improving information retrieval using medical subject headings concepts: a test case on rare and chronic diseases. *Journal of the Medical Library Association: JMLA*, 100(3): 176, 2012.

Thierry Declerck. Integration of the thesaurus for the social sciences (thesoz) in an information extraction system. In *Proceedings of the 7th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 90–95, 2013.

Danilo Dessì, Francesco Osborne, Diego Reforgiato Recupero, Davide Buscaldi, and Enrico Motta. Cs-kg: A large-scale knowledge graph of research entities and claims in computer science. In *International Semantic Web Conference (ISWC)*, 2022.

F. Dhombres and O. Bodenreider. Interoperability between phenotypes in research and healthcare terminologies-investigating partial mappings between hpo and snomed ct. *Journal of Biomedical Semantics*, 7(1), 2016. doi: 10.1186/s13326-016-0047-3.

Anastasia Dimou, Miel Vander Sande, Pieter Colpaert, Ruben Verborgh, Erik Mannens, and Rik Van de Walle. Rml: a generic language for integrated rdf mappings of heterogeneous data. In *Ldow*, 2014. URL http://ceur-ws.org/Vol-1184/ldow2014_paper_01.pdf.

Ying Ding, Ronald Rousseau, and Dietmar Wolfram, editors. *Measuring Scholarly Impact*. Springer International Publishing, Cham, 2014. ISBN 978-3-319-10376-1. doi: 10.1007/978-3-319-10377-8.

Warith Eddine Djeddi and Mohamed Tarek Khadir. A novel approach using context-based measure for matching large scale ontologies. In Ladjel Bellatreche and Mukesh K. Mohania, editors, *Data Warehousing and Knowledge Discovery*, pages 320–331, Cham, 2014. Springer International Publishing. ISBN 978-3-319-10160-6.

Edward Dunne and Klaus Hulek. Mathematics subject classification 2020. *Not. Am. Math. Soc*, 67(3): 410–411, 2020.

Halit Erdogan, Esra Erdem, and Olivier Bodenreider. Exploiting umls semantics for checking semantic consistency among umls concepts. *Studies in health technology and informatics*, 160(0 1):749, 2010.

Jung-Wei Fan and Carol Friedman. Combining contextual and lexical features to classify umls concepts. In *AMIA Annual Symposium Proceedings*, volume 2007, page 231. American Medical Informatics Association, 2007.

Said Fathalla, Christoph Lange, and Sören Auer. An upper ontology for modern science branches and related entities. In *European Semantic Web Conference*, pages 436–453. Springer, 2023.

Michael R Genesereth and Nils J Nilsson. *Logical foundations of artificial intelligence*. Morgan Kaufmann, 2012.

Claudio Gnoli, Koraljka Golub, David Haynes, Athena Salaba, Ali Shiri, and Aida Slavic. Library catalog's search interface: Making the most of subject metadata. *KO KNOWLEDGE ORGANIZATION*, 51(3):169–186, 2024.

Thomas R. Gruber. A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2):199–220, 1993. ISSN 1042-8143. doi: https://doi.org/10.1006/knac.1993.1008. URL https://www.sciencedirect.com/science/article/pii/S1042814383710083.

Huanying Gu, George Hripcsak, Yan Chen, C Paul Morrey, Gai Elhanan, James J Cimino, James Geller, and Yehoshua Perl. Evaluation of a umls auditing process of semantic type assignments. In *AMIA Annual Symposium Proceedings*, volume 2007, page 294. American Medical Informatics Association, 2007.

Huanying (Helen) Gu, Yehoshua Perl, Gai Elhanan, Hua Min, Li Zhang, and Yi Peng. Auditing concept categorizations in the umls. *Artificial Intelligence in Medicine*, 31(1):29–44, 2004. ISSN 0933-3657. doi: https://doi.org/10.1016/j.artmed.2004.02.002. URL https://www.sciencedirect.com/science/article/pii/S0933365704000387.

Michael Halper, C Paul Morrey, Yan Chen, Gai Elhanan, George Hripcsak, and Yehoshua Perl. Auditing hierarchical cycles to locate other inconsistencies in the umls. In *AMIA Annual Symposium Proceedings*, volume 2011, page 529. American Medical Informatics Association, 2011.

Kanyao Han, Pingjing Yang, Shubhanshu Mishra, and Jana Diesner. Wikicssh: Extracting computer science subject headings from wikipedia. In Ladjel Bellatreche, Mária Bieliková, Omar Boussaïd, Barbara Catania, Jérôme Darmont, Elena Demidova, Fabien Duchateau, Mark Hall, Tanja Merčun, Boris Novikov, Christos Papatheodorou, Thomas Risse, Oscar Romero, Lucile Sautot, Guilaine Talens, Robert Wrembel, and Maja Žumer, editors, *ADBIS, TPDL and EDA 2020 Common Workshops and Doctoral Consortium*, pages 207–218, Cham, 2020. Springer International Publishing. ISBN 978-3-030-55814-7.

Heather Hedden. Taxonomies and controlled vocabularies best practices for metadata. *Journal of Digital Asset Management*, 6:279–284, 2010.

Gail M Hodge. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. Number 91. Digital Library Federation, 2000.

Jiaxin Huang, Yiqing Xie, Yu Meng, Yunyi Zhang, and Jiawei Han. 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*, pages 1928–1936, 2020.

Ekaterina V Ilgisonis, Mikhail A Pyatnitskiy, Svetlana N Tarbeeva, Artem A Aldushin, and Elena A Ponomarenko. How to catch trends using mesh terms analysis? *Scientometrics*, 127(4):1953–1967, 2022.

Ernesto Jiménez-Ruiz, Bernardo Cuenca Grau, Ian Horrocks, and Rafael Berlanga. Logic-based assessment of the compatibility of umls ontology sources. *Journal of biomedical semantics*, 2(1):1–16, 2011.

Xia Jing. The unified medical language system at 30 years and how it is used and published: Systematic review and content analysis. *JMIR Med Inform*, 9(8):e20675, Aug 2021. ISSN 2291-9694. doi: 10.2196/20675. URL https://medinform.jmir.org/2021/8/e20675.

Helen L Johnson, K Bretonnel Cohen, and Lawrence Hunter. A fault model for ontology mapping, alignment, and linking systems. In *Biocomputing 2007*, pages 233–244. World Scientific, 2007.

Yannis Kalfoglou and Marco Schorlemmer. Ontology mapping: the state of the art. *The knowledge engineering review*, 18(1):1–31, 2003.

Bharath Kandimalla, Shaurya Rohatgi, Jian Wu, and C. Lee Giles. Large scale subject category classification of scholarly papers with deep attentive neural networks. *Frontiers in Research Metrics and Analytics*, 5:31, 2021. ISSN 2504-0537. doi: 10.3389/frma.2020.600382. URL https://www.frontiersin.org/article/10.3389/frma.2020.600382.

Yong-Bin Kang, Yuan-Fang Li, and Ross L. Coppel. Capturing researcher expertise through mesh classification. In *Proceedings of the 8th International Conference on Knowledge Capture*, K-CAP '15, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450338493. doi: 10.1145/2815833.2815837. URL https://doi-org.libezproxy.open.ac.uk/10.1145/2815833.2815837.

Elisa F Kendall and Deborah L McGuinness. *Ontology engineering*. Morgan & Claypool Publishers, 2019.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. Large language models are zero-shot reasoners, 2023. URL https://arxiv.org/abs/2205.11916.

Sándor Kopácsi, Rastislav Hudak, and Raman Ganguly. 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, Sep. 2017. doi: 10.31263/voebm.v70i2.1897. URL https://journals.univie.ac.at/index.php/voebm/article/view/2075.

Sotiris Kotitsas, Dimitris Pappas, Natalia Manola, and Haris Papageorgiou. Scinobo: a novel system classifying scholarly communication in a dynamically constructed hierarchical field-of-science taxonomy. *Frontiers in Research Metrics and Analytics*, 8, 2023. ISSN 2504-0537. doi: 10.3389/frma.2023.1149834. URL https://www.frontiersin.org/journals/research-metrics-and-analytics/articles/10.3389/frma.2023.1149834.

Thomas S. Kuhn. *The structure of Scientific Revolutions*. The University of Chicago Press, 1962.

P. Lambe. *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness*. Chandos Knowledge Management. Elsevier Science, 2014. ISBN 9781780632001. URL https://books.google.co.uk/books?id=z1mpAgAAQBAJ.

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

F. Löffler, V. Wesp, S. Babalou, P. Kahn, R. Lachmann, B. Sateli, R. Witte, and B. König-Ries. Scholarlensviz: A visualization framework for transparency in semantic user profiles. In Kerry Taylor, Rafael Gonçalves, Freddy Lecue, and Jun Yan, editors, *Proceedings of the ISWC 2020 Demos and Industry Tracks: From Novel Ideas to Industrial Practice co-located with 19th International Semantic Web Conference (ISWC 2020), Globally online, November 1-6, 2020 (UTC).*, 2020.

Yinquan Lu, Wenhao Zhu, Lei Li, Yu Qiao, and Fei Yuan. Llamax: Scaling linguistic horizons of llm by enhancing translation capabilities beyond 100 languages. *arXiv preprint arXiv:2407.05975*, 2024.

Florian Mai, Lukas Galke, and Ansgar Scherp. 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*, JCDL '18, page 169–178, New York, NY, USA, 2018. Association for Computing Machinery. ISBN 9781450351782. doi: 10.1145/3197026.3197039. URL https://doi.org/10.1145/3197026.3197039.

M. Mercedes Martínez-González and María-Luisa Alvite-Díez. Thesauri and semantic web: Discussion of the evolution of thesauri toward their integration with the semantic web. *IEEE Access*, 7:153151–153170, 2019. doi: 10.1109/ACCESS.2019.2948028.

Fulvio Mazzocchi. Knowledge Organization System (KOS): An Introductory Critical Account. *KO KNOWLEDGE ORGANIZATION*, 45(1):54–78, 2018. ISSN 0943-7444. doi: 10.5771/0943-7444-2018-1-54.

Bridget T McInnes, Ted Pedersen, and John Carlis. Using umls concept unique identifiers (cuis) for word sense disambiguation in the biomedical domain. In *AMIA annual symposium proceedings*, volume 2007, page 533. American Medical Informatics Association, 2007.

Joan S Mitchell. Relationships in the dewey decimal classification system. *Relationships in the organization of knowledge*, pages 211–226, 2001.

Elena Montiel-Ponsoda, G Aguado De Cea, Asunción Gómez-Pérez, and Wim Peters. Enriching ontologies with multilingual information. *Natural language engineering*, 17(3):283–309, 2011.

C. Paul Morrey, James Geller, Michael Halper, and Yehoshua Perl. The neighborhood auditing tool: A hybrid interface for auditing the umls. *Journal of Biomedical Informatics*, 42(3):468–489, 2009. ISSN 1532-0464. doi: https://doi.org/10.1016/j.jbi.2009.01.006. URL https://www.sciencedirect.com/science/article/pii/S1532046409000197. Auditing of Terminologies.

E. Motta. *Reusable Components for Knowledge Modelling: Case Studies in Parametric Design Problem Solving*. Frontiers in artificial intelligence and applications. IOS Press, 1999. ISBN 9784274903151. URL https://books.google.co.uk/books?id=S6g8FTcQAVYC.

Fleur Mougin and Olivier Bodenreider. Approaches to eliminating cycles in the umls metathesaurus: Naïve vs. formal. In *AMIA Annual Symposium Proceedings*, volume 2005, page 550. American Medical Informatics Association, 2005.

Xiangming Mu, Kun Lu, and Hohyon Ryu. Explicitly integrating mesh thesaurus help into health information retrieval systems: An empirical user study. *Information Processing & Management*, 50(1): 24–40, 2014.

M. A. Musen. The Protégé Project: A Look Back and a Look Forward. *AI Matters*, 1(4):4–12, Jun 2015.

David Newman, Youn Noh, Edmund Talley, Sarvnaz Karimi, and Timothy Baldwin. Evaluating topic models for digital libraries. In *Proceedings of the 10th annual joint conference on Digital libraries*, pages 215–224, 2010.

OpenAlex. Openalex: End-to-end process for topic classification. 2024.

Francesco Osborne and Enrico Motta. Klink-2: integrating multiple web sources to generate semantic topic networks. In *The Semantic Web-ISWC 2015: 14th International Semantic Web Conference, Bethlehem, PA, USA, October 11-15, 2015, Proceedings, Part I 14*, pages 408–424. Springer, 2015.

Francesco Osborne, Enrico Motta, and Paul Mulholland. Exploring scholarly data with rexplore. In Harith Alani, Lalana Kagal, Achille Fokoue, Paul Groth, Chris Biemann, Josiane Xavier Parreira, Lora Aroyo, Natasha Noy, Chris Welty, and Krzysztof Janowicz, editors, *The Semantic Web – ISWC 2013*, pages 460–477, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

Francesco Osborne, Henry Muccini, Patricia Lago, and Enrico Motta. Reducing the effort for systematic reviews in software engineering. *Data Science*, 2(1-2):311–340, 2019.

María-Antonia Ovalle-Perandones, Juan Gorraiz, Martin Wieland, Christian Gumpenberger, and Carlos Olmeda-Gómez. The influence of european framework programmes on scientific collaboration in nanotechnology. *Scientometrics*, 97:59–74, 2013.

Domenico M Pisanelli, Aldo Gangemi, Massimo Battaglia, and Carola Catenacci. Coping with medical polysemy in the semantic web: the role of ontologies. In *MEDINFO 2004*, pages 416–419. IOS Press, 2004.

Alessia Pisu, Livio Pompianu, Angelo Salatino, Francesco Osborne, Daniele Riboni, Enrico Motta, and Diego Reforgiato Recupero. Leveraging language models for generating ontologies of research topics. In *Text2KG @ ESWC 2024*. CEUR (in press), 2024.

Kenneth H. Plate. Library classification for environmental science. *Journal (Water Pollution Control Federation)*, 38(4):580–584, 1966. ISSN 00431303. URL http://www.jstor.org/stable/25035533.

Raminta Pranckutė. Web of science (wos) and scopus: The titans of bibliographic information in today's academic world. *Publications*, 9(1), 2021. ISSN 2304-6775. doi: 10.3390/publications9010012. URL https://www.mdpi.com/2304-6775/9/1/12.

Junping Qiu, Rongying Zhao, Siluo Yang, and Ke Dong. *Informetrics*. Springer Singapore, Singapore, 2017. ISBN 978-981-10-4031-3. doi: 10.1007/978-981-10-4032-0.

Joe Raad and Christophe Cruz. A survey on ontology evaluation methods. In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management - KEOD, (IC3K 2015)*, pages 179–186. INSTICC, SciTePress, 2015. ISBN 978-989-758-158-8. doi: 10.5220/0005591001790186.

Randolph FR Rasch. The nature of taxonomy. *Image: Journal of Nursing Scholarship*, 19(3):147–149, 1987.

Artem Revenko, Anna Breit, Salma Mahmoud, Mark Szabo, and Tomas Knap. Investigate the impact of contextual information on llms for taxonomy expansion. In *Knowledge Graphs in the Time of Language Models and Neuro-Symbolic AI*. IOS Press, 2024.

David Reymond. Patents information for humanities research: Could there be something? *Iberoamerican Journal of Science Measurement and Communication*, 1(1):6, 2020.

Angelo Salatino, Francesco Osborne, and Enrico Motta. Cso classifier 3.0: a scalable unsupervised method for classifying documents in terms of research topics. *International Journal on Digital Libraries*, 2021. doi: 10.1007/s00799-021-00305-y. URL https://doi.org/10.1007/s00799-021-00305-y.

Angelo Salatino, Simone Angioni, Francesco Osborne, Diego Reforgiato Recupero, and Enrico Motta. Diversity of expertise is key to scientific impact: a large-scale analysis in the field of computer science. 2023. doi: 10.55835/6442f3fd947802668eee976c. URL https://dapp.orvium.io/deposits/6442f3fd947802668eee976c/view.

Angelo A Salatino, Francesco Osborne, and Enrico Motta. Augur: forecasting the emergence of new research topics. In *Proceedings of the 18th ACM/IEEE on joint conference on digital libraries*, pages 303–312, 2018a.

Angelo A. Salatino, Thiviyan Thanapalasingam, Andrea Mannocci, Francesco Osborne, and Enrico Motta. The computer science ontology: A large-scale taxonomy of research areas. In Denny Vrandečić, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl, editors, *The Semantic Web – ISWC 2018*, pages 187–205, Cham, 2018b. Springer International Publishing. ISBN 978-3-030-00668-6.

Angelo A. Salatino, Francesco Osborne, Aliaksandr Birukou, and Enrico Motta. Improving editorial workflow and metadata quality at springer nature. In *The Semantic Web – ISWC 2019*, pages 507–525, Cham, 2019. Springer International Publishing. ISBN 978-3-030-30796-7.

Angelo A. Salatino, Thiviyan Thanapalasingam, Andrea Mannocci, Aliaksandr Birukou, Francesco Osborne, and Enrico Motta. The Computer Science Ontology: A Comprehensive Automatically-Generated Taxonomy of Research Areas. *Data Intelligence*, 2(3):379–416, 07 2020. ISSN 2641-435X. doi: 10.1162/dint_a_00055. URL https://doi.org/10.1162/dint_a_00055.

Angelo Antonio Salatino. Early detection of research trends, 2019. URL http://oro.open.ac.uk/67224/.

Manuel Salvadores, Paul R Alexander, Mark A Musen, and Natalya F Noy. Bioportal as a dataset of linked biomedical ontologies and terminologies in rdf. *Semantic web*, 4(3):277–284, 2013.

Guergana K Savova, James J Masanz, Philip V Ogren, Jiaping Zheng, Sunghwan Sohn, Karin C Kipper-Schuler, and Christopher G Chute. 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, 09 2010. ISSN 1067-5027. doi: 10.1136/jamia.2009.001560. URL https://doi.org/10.1136/jamia.2009.001560.

Thomas Schandl and Andreas Blumauer. Poolparty: Skos thesaurus management utilizing linked data. In Lora Aroyo, Grigoris Antoniou, Eero Hyvönen, Annette ten Teije, Heiner Stuckenschmidt, Liliana Cabral, and Tania Tudorache, editors, *The Semantic Web: Research and Applications*, pages 421–425, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg. ISBN 978-3-642-13489-0.

Andrea Scharnhorst, Katy Börner, and Peter van den Besselaar, editors. *Models of Science Dynamics*, volume 53. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012. ISBN 978-3-642-23067-7. doi: 10.1007/978-3-642-23068-4.

Jingbo Shang, Xinyang Zhang, Liyuan Liu, Sha Li, and Jiawei Han. Nettaxo: Automated topic taxonomy construction from text-rich network. In *Proceedings of the web conference 2020*, pages 1908–1919, 2020.

Hui Shi, Kurt Maly, Steven Zeil, and Mohammad Zubair. Comparison of ontology reasoning systems using custom rules. In *Proceedings of the international conference on web intelligence, mining and semantics*, pages 1–9, 2011.

Ali Asghar Shiri, Crawford W Revie, and Gobinda Chowdhury. Thesaurus-assisted search term selection and query expansion: a review of user-centred studies. *Knowledge organization*, 29(1):1–19, 2002.

Pavel Shvaiko and Jérôme Euzenat. Ten challenges for ontology matching. In Robert Meersman and Zahir Tari, editors, *On the Move to Meaningful Internet Systems: OTM 2008*, pages 1164–1182, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. ISBN 978-3-540-88873-4.

Arnab Sinha, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-June (Paul) Hsu, and Kuansan Wang. An overview of microsoft academic service (mas) and applications. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15 Companion, page 243–246, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450334730. doi: 10.1145/2740908.2742839. URL https://doi.org/10.1145/2740908.2742839.

Peter Sjögårde and Fereshteh Didegah. The association between topic growth and citation impact of research publications. *Scientometrics*, 127(4):1903–1921, 2022.

Peter Sjögårde and Per Ahlgren. Granularity of algorithmically constructed publication-level classifications of research publications: Identification of specialties. *Quantitative Science Studies*, 1(1):207–238, 02 2020. ISSN 2641-3337. doi: 10.1162/qss_a_00004. URL https://doi.org/10.1162/qss_a_00004.

Luke T Slater, Georgios V Gkoutos, and Robert Hoehndorf. Towards semantic interoperability: finding and repairing hidden contradictions in biomedical ontologies. *BMC Medical Informatics and Decision Making*, 20(10):311, 2020. ISSN 1472-6947. doi: 10.1186/s12911-020-01336-2. URL https://doi.org/10.1186/s12911-020-01336-2.

Luca Soldaini and Nazli Goharian. Quickumls: a fast, unsupervised approach for medical concept extraction. In *MedIR workshop, sigir*, pages 1–4, 2016.

Alessandro Solimando, Ernesto Jiménez-Ruiz, and Giovanna Guerrini. Detecting and correcting conservativity principle violations in ontology-to-ontology mappings. In Peter Mika, Tania Tudorache, Abraham Bernstein, Chris Welty, Craig Knoblock, Denny Vrandečić, Paul Groth, Natasha Noy, Krzysztof Janowicz, and Carole Goble, editors, *The Semantic Web – ISWC 2014*, pages 1–16, Cham, 2014. Springer International Publishing. ISBN 978-3-319-11915-1.

Juliane Stiller, Violeta Trkulja, Sophie Biesenbender, and Vivien Petras. 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, 2021. URL http://web.archive.org/web/20211004184211/https://www.kerndatensatz-forschung.de/docs_ff/2021-04-30_ffk-projekt_dokumentation.pdf.

Cassidy R. Sugimoto and Vincent Lariviére. *Measuring Research: What Everyone Needs to Know*, volume 68. Oxford University Press, July 2018.

Ed Summers, Antoine Isaac, Clay Redding, and Dan Krech. Lcsh, skos and linked data. In *Proceedings of the 2008 International Conference on Dublin Core and Metadata Applications*, DCMI '08, page 25–33. Dublin Core Metadata Initiative, 2008.

Thiviyan Thanapalasingam, Francesco Osborne, Aliaksandr Birukou, and Enrico Motta. Ontology-based recommendation of editorial products. In Denny Vrandečić, Kalina Bontcheva, Mari Carmen Suárez-Figueroa, Valentina Presutti, Irene Celino, Marta Sabou, Lucie-Aimée Kaffee, and Elena Simperl, editors, *The Semantic Web – ISWC 2018*, pages 341–358, Cham, 2018. Springer International Publishing. ISBN 978-3-030-00668-6.

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

T. Tudorache, C. Nyulas, N. F. Noy, and M. A. Musen. WebProtégé: A Collaborative Ontology Editor and Knowledge Acquisition Tool for the Web. *Semant Web*, 4(1):89–99, Jan 2013.

Thanasis Vergoulis, Serafeim Chatzopoulos, Theodore Dalamagas, and Christos Tryfonopoulos. Veto: Expert set expansion in academia. In Mark Hall, Tanja Merčun, Thomas Risse, and Fabien Duchateau, editors, *Digital Libraries for Open Knowledge*, pages 48–61, Cham, 2020. Springer International Publishing. ISBN 978-3-030-54956-5. doi: 10.1007/978-3-030-54956-5_4.

Denny Vrandečić and Markus Krötzsch. *Semantic MediaWiki*, pages 171–179. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009. ISBN 978-3-540-88845-1. doi: 10.1007/978-3-540-88845-1_13. URL https://doi.org/10.1007/978-3-540-88845-1_13.

Dominic Widdows, Stanley Peters, Scott Cederberg, Chiu-Ki Chan, Diana Steffen, and Paul Buitelaar. 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*, pages 9–16, 2003.

Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, Anders Wesslén, et al. *Experimentation in software engineering*, volume 236. Springer, 2012.

Erjia Yan. Research dynamics: Measuring the continuity and popularity of research topics. *Journal of Informetrics*, 8(1):98–110, 2014.

Heyoung Yang and Hyuck Jai Lee. Research trend visualization by mesh terms from pubmed. *International Journal of Environmental Research and Public Health*, 15(6), 2018. ISSN 1660-4601. doi: 10.3390/ijerph15061113. URL https://www.mdpi.com/1660-4601/15/6/1113.

Hong Yung Yip, Vinh Nguyen, and Olivier Bodenreider. Construction of umls metathesaurus with knowledge-infused deep learning. In *BlockSW/CKG@ ISWC*, 2019.

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

Benjamin Zapilko, Johann Schaible, Philipp Mayr, and Brigitte Mathiak. Thesoz: A skos representation of the thesaurus for the social sciences. *Semant. Web*, 4(3):257–263, July 2013. ISSN 1570-0844.

Marcia Lei Zeng. Knowledge organization systems (kos). *KO KNOWLEDGE ORGANIZATION*, 35(2-3):160–182, 2008.

Marcia Lei Zeng and Philipp Mayr. Knowledge organization systems (kos) in the semantic web: a multi-dimensional review. *International Journal on Digital Libraries*, 20(3):209–230, 2019.

Chao Zhang, Fangbo Tao, Xiusi Chen, Jiaming Shen, Meng Jiang, Brian Sadler, Michelle Vanni, and Jiawei Han. 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*, pages 2701–2709, 2018.

Xiaoyu Zhang, Senthil Chandrasegaran, and Kwan-Liu Ma. Conceptscope: Organizing and visualizing knowledge in documents based on domain ontology. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2021.