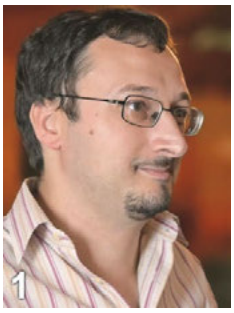
Library Catalog's Search Interface: Making the Most of Subject Metadata

Claudio Gnoli*, Koraljka Golub**, David Haynes***,
Athena Salaba****, Ali Shiri*****, and Aida Slavic******

* University of Pavia, Science and Technology Library – Pavia, Italy
** Linnaeus University, Department of Library and Information Science, School of Cultural Sciences, Faculty of Arts and Humanities – Kalmar and Vaxjo, Sweden
*** Edinburgh Napier University – Edinburgh, United Kingdom of Great Britain and Northern Ireland
**** Kent State University, School of Information – Kent, United States
***** University of Alberta, School of Library and Information Studies – Alberta, Canada
****** UDC Consortium – Hague, Netherlands and University of Zagreb, Department of Information and Communication Sciences, Zagreb – Croatia
* claudio.gnoli@unipv.it, ** koraljka.golub@lnu.se, *** d.haynes@napier.ac.uk,
**** asalaba@kent.edu, ***** ashiri@ualberta.ca, ****** aida.slavic@udcc.org

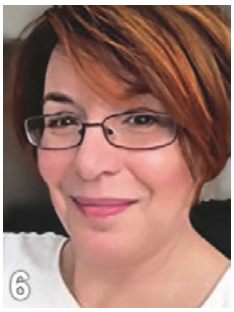
Claudio Gnoli is librarian at the University of Pavia Science and Technology Library. His interests include knowledge organization in general, theory of classification, their philosophical grounds and their application to online knowledge bases. He has authored papers and books in the KO field, including Introduction to Knowledge Organization (Facet, 2020), and is a member of the editorial boards of Journal of Documentation, Knowledge Organization and AIDA Informazioni. He is a former vice-president and the current webmaster of international ISKO. His website is at gnoli.eu.



Koraljka Golub is a full professor and head of iInstitute at Linnaeus University, Sweden. Her research primarily focuses on topics related to information retrieval and knowledge organization. Of her particular interest is the integration of traditional knowledge organization systems with social tagging and/or automated subject indexing and evaluating results in the context of end-user information retrieval. Details of her research projects and related activities are available at her website koraljka.net/.

David Haynes teaches knowledge organization at Edinburgh Napier University. He is Secretary of ISKO and chairs the ISKO Working Group on Metadata Guidelines for Subject Access. He gained a PhD in Information Science from City University of London after a 30-year career as an LIS researcher, consultant and project manager. He is the author of the book: Metadata for Information Management and Retrieval (2nd edition, 2018).

Athena Salaba is a Professor in the School of Information at Kent State University, USA. Her teaching and research areas include organization of information, knowledge organization systems, metadata, subject access, presentation of metadata through system interfaces, and user interactions with bibliographic information and interfaces. She has published books, peer-reviewed articles and papers, presented on these topics internationally, and has been involved in the development and review of the FRSAD and LRM bibliographic conceptual models. She served as the Chair of the IFLA Subject Analysis and Access Section and serves as the Treasurer for the International Society for Knowledge Organization (ISKO).



https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Ali Shiri is a Professor in the School of Library and Information Studies at the University of Alberta, Edmonton, Alberta, Canada, and is currently the Vice Dean of the Faculty of Graduate and Postdoctoral Studies. He received his PhD in Information Science from the University of Strathclyde Department of Computer and Information Sciences in Glasgow, Scotland in 2004. Ali has been teaching, researching, and writing about digital libraries and digital information interaction, knowledge organization, data and learning analytics, and more recently on AI and ethics in education. Ali's research on cultural heritage digital libraries has been funded by the Social Sciences and Humanities Research Council of Canada (SSHRC).

Aida Slavic the editor-in-chief of the Universal Decimal Classification (UDC) and works on the development and the maintenance of the scheme for the UDC Consortium, based in The Hague. She is also an Associate Professor (visiting lecturer) in the Department of Information and Communication Sciences at the University of Zagreb (Croatia), where she has been lecturer at the master and doctoral programmes on the subject of indexing languages, knowledge organization, information retrieval and classification systems. She holds a Ph.D. in library and information studies from University College London. Her research interest is in use of classification in the networked environment, knowledge organization, classification, metadata and semantic technologies.

Gnoli, Claudio, Koraljka Golub, David Haynes, Athena Salaba, Ali Shiri, and Aida Slavic. 2024. "Library Catalog's Search Interface: Making the Most of Subject Metadata". *Knowledge Organization* 51 (3): 169-186. 88 references. DOI:10.5771/0943-7444-2024-3-169.

Abstract: This article addresses the underutilization of knowledge organization systems (KOS) elements in online library catalogs, hindering effective subject-based search and discovery. It highlights the International Society for Knowledge Organization's initiative to develop metadata guidelines for library catalog procurement, focusing on maximizing the value of subject metadata from classification systems and controlled vocabularies. The paper discusses the rationale for quality subject access, proposes desirable search functionalities based on research, explores implementation challenges, and outlines future developments. The conclusion emphasizes the importance of providing quality subject access in digital services and calls for further research on interface design, guideline adoption, KOS evolution, and the impact of language models on subject metadata use. The work underscores the need for applying controlled vocabularies in search interfaces across libraries, archives, and museums while acknowledging the complementary role of alternative approaches like social tagging and automatic indexing. Extensive future research is suggested to implement search functionalities, promote guidelines adoption, enhance KOS evolution, and assess the influence of language models on subject metadata utilization.

Received: 20 December 2023; Accepted: 29 December 2023

Keywords: Knowledge organization systems; catalogs; metadada guidelines;subject access

† Derived from the set of metadata guidelines for the procurement of library catalogs developed by the International Society for Knowledge Organization (ISKO) working group on Subject Access Metadata.

1.0 Introduction

Online library catalogs do not typically make effective use of knowledge organization systems (KOS) elements to support search and discovery of content by subjects or topics. Content indexing represents considerable intellectual effort by cataloguers but is often overlooked during the procurement process (and sometimes in system design). This article arises from a concern that library and related information search systems have been 'dumbed down' and that there is limited scope for structured subject searching of resources. Most systems offer a simplified search interface and a 'black box' approach, which automatically modifies searches to retrieve large data sets. Much of the effort is then invested in some ranking system using different algorithms based on a combination of statistical analysis and machine learning to present the results. These solutions often do not use the existing subject metadata created to increase precision and relevance.

In order to help address these challenges, the International Society for Knowledge Organization (ISKO) has set up an international working group to develop a set of metadata guidelines for the procurement of library catalogs (ISKO STAC working group on Subject Access Metadata – https://www.isko.org/stac/metadata). The guidelines aim to ensure that metadata-based search systems, such as those used in libraries, enable users to get maximum value from subject metadata comprising classifications and controlled vocabularies. Although in its initial scope, this proposal is focused on academic libraries and related discovery systems, many aspects of the guidelines will be applicable to other digital library search interfaces/solutions, institutional repositories, as well as information systems of cultural heritage institutions like archives, museums, and galleries.

This paper aims to provide the rationale for offering quality subject access at the interface level and the role that KOSs play in that. Based on earlier research, desirable search functionalities are proposed. The implementation chal-

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

lenges are discussed, which will pave the way for future developments.

The remainder of the paper is structured as follows: the 'Background' section defines key terms, followed by a rationale for ensuring good subject searching and ways to achieve that; the 'Search functionalities in practice' discusses today's most common challenges, with examples from different information search systems; the 'Subject search functionalities' a list of desirable functionalities; and the paper ends with a 'Conclusion,' summarizing key points and presenting guidelines for future research.

2.0 Background

2.1 Terminology

2.1.1 Knowledge organization systems

The term "knowledge organization systems" (KOS) is used to denote terminological and classification systems, tools, and services developed to organize knowledge sources by subject and to present the organized interpretation of knowledge structures, including automated categorization for task-oriented applications or knowledge mining software (Golub 2011). Our focus here is on KOS designed for information discovery.

There are two expressions related to KOSs: *controlled vocabulary* and *indexing language*. These can be considered different types of KOS, each with a different scope. The expression "controlled vocabulary" can denote any controlled list of terms used in metadata to describe documents. The term "Indexing languages" refers to a specific kind of KOS representing formalized languages used to describe the subject content of documents for information retrieval purposes (Golub 2011; Mazzocchi 2018).

The two main types of indexing language are: alphabetical (using a selection of natural language terms, thus requiring terminology control, such as thesauri, descriptor systems, and subject heading lists) and systematic (classifications, mostly using symbols, operating with concepts and not being primarily concerned with natural language). The main characteristics of indexing languages are that they are concerned primarily with the subject content of documents and that they contain rules for applying them and, in some cases, syntax rules for pre-combination of terms in the process of indexing.

The general purpose of a KOS is to provide a means for organizing information (ANSI/NISO Z39.19 [3]), through:

- translation of the natural language of authors, indexers, and users into a vocabulary that can be used for indexing and retrieval;
- ensuring consistency through uniformity in term format and in the assignment of terms indicating semantic relationships among terms; and,
- supporting browsing by providing consistent and clear hierarchies in a navigation system supporting retrieval.

KOSs play a crucial role in resource retrieval and discovery. They improve the effectiveness of retrieval by helping to handle the sheer mass of available information. They also provide knowledge-based support for end users who access information without the help of an intermediary. In comparison to free-text searching, there are many advantages to searching by KOS terms:

- the most relevant search terms are selected, and relevant search terms that are not explicitly mentioned in a document may be added;
- search terms are controlled, i.e. disambiguated, so that there is no confusion between terms that look the same but have different meanings; and,
- search terms can come from semantically structured vocabularies – hence documents can be found by searching for synonyms, narrower, broader, and even conceptually related terms that may not be present in the document itself (semantic query expansion).

A well-structured KOS can be used as the knowledge base for an interface that can assist users with search topic clarification (e.g., through browsing well-structured hierarchies and guided facet analysis) and with finding good search terms (e.g., through query term mapping and query term expansion: synonyms and hierarchical inclusion).

2.1.2 Information discovery services

A recent trend of online library catalogs is that they may also provide access to books and journals available to library users in digital format, as well as resources from outside the library like e-books, journal articles from commercial databases, pre-prints; as such, they have been referred to as integrated catalogs or web-scale discovery services. Discovery layers, discovery interfaces and discovery tools are also common terms. In this work, the terms integrated catalog and discovery *service* are used depending on the context of the author or topic discussed.

3.0 Challenges of subject searching

While support for subject searching has been traditionally advocated for in library catalogs, notably since Cutter's objectives for library catalogs (1876), research shows that since the library automation from the 1980s onwards and throughout all generations of library systems, subject access

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

in online library catalogs has not been satisfactory (Markey 2007; Dempsey 2012; Golub 2016). Library catalogs are part of library management systems (LMS), which can take different forms, including Web-based discovery services, which serve as one-stop-for-all resources to which the library has access. More recent developments and adoptions of such systems try to match users' expectations by implementing Google-like single search box interfaces. However, as it is not possible to apply efficient web search engine ranking mechanisms and the exploitation of intellectual effort invested into subject indexing and classification is missing from these services, retrieval failures are common.

Subject searching is an important requirement in online search systems such as library catalogs (Hider and Liu 2013; Hunter 1991; Villen-Rueda et al. 2007; Wells 2020), bibliographic databases (Siegfried et al. 1993), repositories (Heery et al. 2006), discovery services (Meadow and Meadow 2012), online museums (Baca 2004; Liew 2004), and related digital search services (Patel et al. 2005).

However, in comparison to known-item searching (e.g., queries for information objects whose title, author, etc. are known beforehand), searching by subject, even if all resources are available in digital form, often proves much more challenging. This may be due to:

1. The difficulties of formulating queries with insufficient knowledge of the subject matter. One search box does not help the user see what information resources are available beyond the search box. Instead, an overview of some kind would be useful (see, e.g., Gnoli and Cheti 2013; Ackerhurst and Polvere 2020), and so would help the choice of the right search term or class (e.g. by browsing a tree of concepts from a classification scheme; an example is University of Pavia's SciGator – http://scigator.unipv.it/indexe.php) (Lardera et al. 2017).
2. Insufficient knowledge of the resources covered by the information system, resulting in an inability to use right search terms (Belkin et al., 1982).
3. Insufficient knowledge of information searching (i.e. how to formulate a search query to reflect the information need).
4. Semantic ambiguities inherent to natural language: the same word can take on different meanings (polysemy) that could be completely unrelated (homonymy), while one concept can, in turn, be named using different words (synonymy). Terminological polysemy leads to the retrieval of irrelevant results: in large databases, this may mean too many results to review manually and users typically do not browse beyond the first page with top ten results. For example, an author of a document may use a different word for a certain concept than the user does in their search query. In such cases, automatic retrieval mechanisms would not retrieve that document. Synonymy presents challenges to effective searching by placing the burden on the searcher, who would ideally need to include all possible synonyms in a query and connect them by the Boolean operator OR (which is usually not the default operator) in order to obtain a comprehensive set of results. Homonymy leads to queries that often end up producing false positives. For example, a user may want to find resources about 'bank' related to rivers, while the system may also retrieve 'bank' related to financial institutions.
5. Semantic ambiguities arising from multiple-word search terms.
6. Texts do not always explicitly name concepts that they write about. For example, searching for publications in the field of digital humanities will result in incomplete results because the term may not be used by an author who does not like that term or because the author works with digital archaeology and does not include the broader term *digital humanities* (or even the term *digital archaeology* as it may not been needed to mention).
7. In many humanities disciplines and works of literary fiction, language is often metaphorical on purpose, with related themes being intertwined with blurred boundaries between them. This makes it hard to find the resources hard to find by separate terms. Consider the example of fiction; or, in particular, lesbian, gay, bisexual, transgender, queer, intersex (LGBTQI) fiction – it has been reported that even librarians miss identifying works as LGBTQI unless the themes are obvious from a book cover or a review (De la Tierra 2008, Golub et al. 2022).
8. Texts from different historical periods often use different terms due to lexical and grammatical changes for the same concepts than we do today, and these terms may also be expressed through contested historical language (see Gnoli 2014 on Marc Bloch). Concepts themselves and associated terms can be subject to semantic shifts and changes in meaning and usage.
9. Older texts and manuscripts that have been digitized will also often have misspelled terms due to challenges with optical character recognition (OCR), resulting in not retrieving relevant documents or possibly false positives.
10. The problem is exacerbated with non-textual media that do not lend themselves to full text searching or do not have a narrative and are open to interpretation (Svenonius 1994). Relying on the non-textual content (of e.g., artwork, music, performing arts, intangible cultural heritage) is hard to capture well by an automated IR approach and may be challenging even for an experienced subject indexing professional. The representa-

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

tion of 'ofness', 'isness', and 'aboutness' is often needed to index non-textual materials. Panofsky describes three levels of meanings for works of art, requiring pre-iconographic, iconographic, and iconological analysis (Panofsky 1939). Based on Panofsky's theory, Shatford defines four subject facets of "who," "what," "where," and "when" with three aspects: the "generic of," the "specific of," and the "about." Shatford's faceted classification provides a structure for systematically identifying possible subjects (Shatford 1986, 1994). It is most likely that only very specialized KOS would accommodate such a level of indexing for non-textual materials. Many of these information objects are unique (rather than found in multiple library databases) and the only way to discover them is through associated metadata containing KOSs terms.

4.0 Cataloguing for subject access

In order to alleviate these problems, online search services should enable the use of assigned subject indexing, a process in which subject terms are taken from controlled vocabularies such as subject headings systems, thesauri, and classification systems. These are designed to help the user select a more specific concept to increase precision, a broader concept or related concepts to increase recall, to help the user disambiguate between homonyms, to discover which term is best used to name a concept or browse through a category of items on the same subject. In addition, hierarchical browsing of classification schemes and other systems with hierarchical structures could help the user improve their understanding of their information requirements and formulate their queries more accurately.

ISO 5963:1985, which was confirmed in 2020 (International Organization for Standardization 1985), prescribes general techniques for subject indexing and clearly states that these are to be applied "by any agency in which human indexers analyze the subjects of documents and express these subjects in indexing terms" (2020, 1), defining documents to be "any item amenable to cataloging or indexing, specifically including non-print media and three-dimensional objects or realia". The standard gives a document-oriented definition of manual subject indexing as a process involving three steps: (1) determining the subject content of a document; (2) a conceptual analysis to decide which aspects of the content should be represented; (3) translation of those concepts or aspects into controlled vocabulary terms or notations.

Objectives of library catalogs in relation to subject access have been traditionally anchored in Cutter's 'objects', as he called them, which are to: 1) enable finding an item of which the subject is known; 2) show what the library has on a given subject; and 3) assist in the choice of a book as to its topical character (Cutter 1876, 5). These objects have been an integral part of cataloging codes ever since and continue to be so in the contemporary FRBR (Functional Requirements for Bibliographic Records) family of conceptual models for catalog functionality. The FRBR family includes:

- *Functional Requirements for Bibliographic Records (FRBR)*;
- *Functional Requirements for Authority Data (FRAD)*; and,
- *Functional Requirements for Subject Authority Data (FRSAD)*.

In 2017, these three models were consolidated into the IFLA Library Reference Model (Riva et al. 2017). The consolidated model prescribes five user tasks, which need to be translated into cataloging rules to account for relationships between works, expressions, manifestations, and items and relationships between topics and these works, expressions, manifestations, and items. In the context of subject access, *IFLA LRM* and *FRSAD* (Zeng et al. 2011) tasks of finding, identifying, selecting, obtaining, and exploring, could be applied as:

- Find: to find resources embodying works that are described by a given subject label, for example, search using a term or symbol (nomen) that is used in a subject headings system or a classification scheme;
- Identify: to clearly understand the nature of the resources found and to distinguish between similar resources, e.g., those that are indexed by homonyms, or those with the same topic but in a different context i.e. the same concept may be studied in different fields of knowledge and from different perspectives (e.g., virus in medicine, virus in biology, or virus in public health);
- Select: to determine the suitability of the resources found and to choose (by accepting or by rejecting) specific resources that seem the most relevant, e.g., due to certain aspects, characteristics or approach to the subject described;
- Obtain: to access the content of the resource;
- Explore: to use the relationships between one resource and another, to place them in a context, e.g., to browse around related topics such as through using related terms in a thesaurus, or to see narrower and broader terms or classes, in order to understand the relationships between various nomens for an entity such as: examine the variant names for a subject within a controlled vocabulary, survey the variant terms used in different contexts of use, which may include different languages; explore correlations between nomens for the same entity in different

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

controlled vocabularies, e.g., finding a thesaurus descriptor which corresponds to a classification number.

Typically, bibliographic description standards, such as the *International Standard Bibliographic Description* (ISBD) or national cataloging rules (e.g., *AACR2 Anglo-American Cataloging Rules*), were only concerned with the formal characteristics of the documents and did not mention subject cataloging. The issues of subject description were dealt with within indexing and classification guidelines and textbooks. The recent international standard for resource description that follows an object-oriented approach to resource metadata, *Resource Description and Access* (RDA), makes an effort to point out that subject representation or relationship to the subject of a work is needed: "The RDA element for the subject relationship generally reflects the relationship associated with the entity work as defined in FRSAD" (Kuhagen 2015, 3). Section 7 of the original RDA covers the relationships used to find works on a particular subject. Chapter 23 is titled "General Guidelines on Recording Relationships Between Works and Subjects" (RDA Co-Publishers 2010). It defines the relationship element "subject" as "a topic that a work is about". However, this has not been extensively elaborated beyond the definitions of the subject relation, the subject relationship element and its subtypes, and the guidance to use KOS for the element values; and therefore, concrete guidelines for the practice of subject indexing are lacking.

5.0 Search functionalities in practice

5.1 General problems

In spite of over a hundred years of developing and implementing bibliographic standards, including those related to subject indexing, bibliographic systems have never reached their full potential in supporting subject searching. Applications of information retrieval technologies have been dominating and poorly related to research in knowledge organization (Hjørland 2021). As a result, by the time the World Wide Web became prevalent, the demand to implement functionalities similar to global search engines such as Google and other commercial services like Amazon, was increasing. These included the single search box, attractive web design, relevance ranking of results, recommendations, and access to a wide range of resources. Although seemingly attractive, these requirements come with problems related to searching based almost entirely on full-text indexing (cf. above), which means that each search would result in millions of hits with no guarantee that the top-ranked ones will address your desired topic in depth or at your level of understanding (Markey 2007). "Faceted" navigation has become a standard feature in large bibliographic systems and subjects are often seen as only one of the "facets" (Chickering and Yang 2014). However, studies show that this can be confusing to end users who frequently do not understand how such facets work or know the types of terms included in them (Emanuel 2011; Osborne and Cox 2015). On the other hand, examples of interfaces that leverage faceted classifications or thesauri, such as Colon Classification or Art and Architecture Thesaurus (Tudhope and Binding 2008; Broughton 2023), are unfortunately uncommon.

The most common issues regarding subject searching today are inconsistency and incompleteness of metadata and the blending of controlled vocabularies, free keywords and full- text automatic indexing (Dempsey 2012; Fagan 2011; Golub, 2016). Commercial bibliographic indexing and abstracting services claim to provide comprehensive coverage, yet their indexing policies are not standardized. Neither are there any international standards or common guidelines for implementing quality-controlled subject access in repositories; instead, a range of different vocabularies are applied across repositories and repository platforms (see, for example, Bundza, 2014).

Today, although both FR family of conceptual models and RDA description guidelines have emphasized the subject and the end user, these aspects remain insufficiently researched or supported by implementation and application guidelines (Cossham 2013) and applied in practice. The official RDA, released in December 2020 but not widely implemented yet, takes a step further and defines the work attribute element "subject" for use to record the topic of a work. Twelve element subtypes have been defined to record the subject relationship. The IRI for the "subject" element can be accessed at http://rdaregistry.info/Elements/w/P10256.

Research has shown that KOSs are particularly needed in large databases covering broad areas of knowledge (Markey 2007; Tibbo 1994) as well as in databases of primary sources (Bair and Carlson 2008) such as museum objects, which cannot be queried using full-text searches alone. Tibbo (1994) makes the point that the exponentially increasing volume of information objects available online leads to information overload and entropy, rather than increasing benefit from access to information. Although full-text searching works for some tasks, for others it creates information overload and prevents the searcher from gaining a comprehensive overview of a topic: if a query returns thousands of retrieved documents, few searchers will browse beyond the first dozen or two hits.

According to East (2007), who studied subject retrieval from ten full-text databases, the databases do not meet the needs of users. East concludes that in the "first phase of digitization", the excitement generated by the prospect of technology to make everything available at a few keystrokes, "has blinded many commentators to the distinct limitations"

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

(239). He claims that the users will use these collections only "to look for documents that they have identified from other sources rather than performing subject searches to discover further resources" (239) East also expresses a wish for a second phase of digitization, "in which the 'quick and dirty' digital libraries of today will be enriched and enhanced to become resources that can effectively meet the information needs of scholars" (239). The availability of funding, or lack thereof, is likely a root cause behind the differences in coverage and effectiveness of indexing between science and humanities, the latter having much less commercial relevance.

To counter high recall with hundreds or thousands of hits and low precision, specific subject indexing should be implemented, involving (1) indexing policies that promote a high level of specificity and (2) indexing languages that are deep and detailed for any given topic, especially for large databases and cross-search services with tens of millions of records. The indexing language needs to have extensive coverage and a rich network of semantic relationships in order to account for the fact that any topic can appear in many different contexts, and topics may be addressed from a very wide range of different perspectives, as emphasized by several researchers in the León Manifesto (2007). Furthermore, specific domains may require their own specialized indexing languages rather than a one-size-fits-all approach (Tibbo 1994), which then also requires a meta subject indexing language, usually called a 'switching language,' that brings them all together in order to support searching across disciplines in an interoperable manner. (although Tibbo focuses on humanities, these apply to any cross-search service). It is worth noting that the idea of a switching language was discussed in numerous projects from the 1960s onwards, one notable example being the *Broad System of Ordering (BSO): a faceted classification* that was designed in the early 1980s to perform the role of mapping between many specific indexing languages and enable cross-collection searching between bibliographic databases. It was published before online systems were commonplace, resulting in it never being implemented or tested in an online environment.

The following sections address the problem in library catalogs, discovery services, as well as other related cross-search services.

5.1.1 Library catalogues

Many researchers have addressed the problematic subject access to information in online library catalogs, pointing to continuing challenges for end users (e.g., Casson et al. 2011 summarizing a wide multi-year survey of Italian catalogs). Barton and Mak (2012) give an overview through a discussion of three generations of online library catalogs (framework set by Hildreth 1984). Key points are briefly presented below.

First-generation online public library catalogs (OPACs) were developed with a focus on efficiency resulting from automation rather than having service to end users in mind. Their functionalities were restricted to exact matching of known-item searches by author, title, or control number; effectively, this was a card catalog in the online form. Second-generation online catalogs supported post-coordinate subject searching using Boolean operators, which, while an improvement in terms of functionalities, proved counterintuitive and hard to use. Third-generation catalogs were developed as experimental systems, e.g., Okapi and Cheshire, and research concluded that their functionalities should include, among others, post-Boolean probabilistic searching, automatic spelling correction, term weighting, relevance feedback, output ranking, and support for finding strategies.

Markey (2007) provides ten reasons why these solutions were not applied to online library catalogs, among them: the failure of library systems' vendors to monitor shifts in information-retrieval technology and respond accordingly with system improvements; the failure of the research community to arrive at a consensus about the most pressing needs for online catalog system improvement; decreasing funding and at the same time the high cost of integrated library systems.

As a result, by the time the World Wide Web became prevalent, OPACs were still second-generation catalogs, and the demand to implement functionalities of global search engines, such as Google and other commercial services like Amazon was increasing. These included a single search box, attractive web design, relevance ranking of results, recommendations, and access to a wide range of resources. However, Markey (2007) argued that the new directions of developments toward simplification would not attract users back to the online catalog. In integrated library catalogs, each search would result in "millions of hits with no guarantee that the top-ranked ones will address your desired topic in depth or at your level of understanding" (Markey 2007).).

Instead, Markey (2007) called for a redesign of an online library catalog that embraces:

1) post-Boolean probabilistic searching on full text;
2) subject cataloging, to help end user define the query, but also improve ranking algorithms by assigning high weights to subject headings, class numbers, as well as back-of-the-book indexes and entries from tables of contents;
3) 'qualification cataloging', as she calls it, i.e., adding metadata like genre, purpose, reviews, academic level, etc., which would allow end users to customize retrieval according to their level of understanding; such metadata

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

could be in part contributed by end users through Web 2.0 functionalities (e.g., folksonomies).

5.1.2 Discovery services

Discovery services today predominantly operate on one integrated index of metadata from all resources involved. A single index provides faster retrieval than distributed searching, which compiles information from different databases on the fly (Barton and Mak 2012). In order for this one central index to operate well, contributing metadata elements and its values need to be interoperable. While metadata are standardized for many uses today, when brought together, they have to be mapped to all other metadata standards used in the integrated index. Furthermore, values such as author names, place names and topics need to be mapped, too. Lastly, metadata policies at different involved institutions need to be harmonized; for example, large research libraries may have subject indexing policies aimed at a greater level of specificity and exhaustivity, than do some more general collections for the general public; the same holds for the choice of metadata elements – different collections may use a different subset of elements from the same metadata standard, or they may implement them with a certain level of difference.

Harmonizing this mix of metadata elements, their values, and indexing policies across collections of resources would ensure that discovery services could fulfill established objectives of a library catalog, ensuring control over search (see above). Ellero (2013), in her analysis of 45 studies of discovery services, concludes that they are "only as effective as the quality and completeness of the metadata they ingest, process, and index...". Indeed, the most common issues regarding subject searching are those of inconsistent and incomplete metadata and blending of controlled vocabularies, free keywords, and full-text automatic indexing (Dempsey 2012; Fagan 2011). Majors (2012) conducted a task-based usability test of five next-generation catalog interfaces and discovery tools with undergraduates across all academic disciplines. Major findings related to subject access show the need to provide context of what has been searched and what is not included. Lee and Chung (2016) studied search effectiveness of discovery services, comparing web-scale discovery services against four individual databases in the fields of Education and Library and Information Science by EBSCO. Based on a small sample of queries and evaluators, it was concluded that the discovery service was less effective than individual databases.

Tarulli (2016) addresses problems of integrating metadata from sources beyond library catalogs and issues that arise from reliance on vendors. A key point emphasized is the need for transparency on how integrated indexes function, particularly regarding ranking and "facet" creation. Yang and Hoffman (2011), who surveyed academic libraries from 260 colleges and universities, showed that the circulation statistics were not part of the algorithm. If Google's success is attributed to ranking based on popularity, it is important for libraries to mimic good ranking, too, and not just the simple-search-box interface. "Faceted" navigation has become a standard feature in discovery tools and subject seems to be often seen as one of the facets (Chickering and Yang 2014), despite their original sense in facet analysis (Broughton, 2023); however, studies point to confusion arising among end users and their lack of understanding of how facets work and the type of terms included in them (Emmanuel 2011; Osborne and Cox 2015).

Prerequisites for harmonization exist to a certain level: many crosswalks of metadata elements, as well as controlled vocabularies, are already available. Furthermore, a significant number of metadata standards and controlled vocabularies with their mappings have made it into linked data and the Semantic Web; see, for example, Library of Congress Linked Data Service, or FAST (Faceted Application of Subject Terminology) which links real-world entities to DBpedia, VIAF and GeoNames.

Therefore, a question arises whether libraries require vendors of discovery services to preserve the established objectives of library catalogs. When selecting a discovery system, Olson (2010) found that libraries often do not approach the decision-making process based on well laid-out arguments for needed features. Instead, reasons for a decision include saving money, facilitating a departmental reorganization, or improving the public perception of the library by implementing something new. A move towards standardization in order to bridge issues preventing unified search is NISO Open Discovery Initiative (ODI) (National Information Standards Organization 2018; Walker 2015). ODI creates a technical recommendation and model for data exchange, which serves as a way for libraries as content providers to work with discovery service vendors. Apart from simplifying the data exchange, it ensures that the vendors follow fair and unbiased indexing and linking practices.

Quality-controlled subject access in examined discovery services seems severely hindered (for an overview, see Golub 2018). This is in spite of the fact that huge resources have been allocated to adding subject index terms from indexing languages to library catalogue records. Little of this is adding value to existing interfaces. While imitating Google's black box approach, the task to retrieve relevant resources to a search query is addressed without making use of the existing index terms, relationships and structures of applied subject indexing languages.

Terms like "Subject", "Keyword", "Category" are loosely and sometimes interchangeably used, but it is not stated anywhere what kind of controlled vocabulary it is, if any, or what the differences are between them. The end user is not

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

informed of the lack of mappings between the different KOSs that are used. This prevents truly integrated cross-searching on a certain subject: using a controlled term from one KOS as a search term will not retrieve relevant documents as the term is not used to describe relevant documents from other systems. Furthermore, there is an obvious loss of the specificity and granularity that KOSs traditionally used by libraries have provided, for example in subject headings.

In literature (Majors 2012; Ellero 2013; Tarulli 2016; Golub 2018), discovery services are criticized for the lack of transparency on the processes behind the scenes, lack of mappings between metadata elements and values thereof, and overwhelming number of results. The fact that results of test searches appear to be complex and confusing is in part due to merging of a number of resource collections, each using different indexing systems. This implies that providing widened search in loosely-controlled discovery services as opposed to traditional OPACs or individual databases of journal articles is not necessarily an advantage.

In terms of LRM and FRSAD, the potential of controlled vocabularies has not been utilized to address the following user tasks:

1. To find, as different resources are indexed using different controlled vocabularies, and also most probably following different indexing policies as they come from different collections of resources;
2. To identify, as homonyms are not disambiguated, different perspectives are not disambiguated, at least not systematically by taking advantage of controlled vocabularies;
3. To select, as aspects, facets or approach to the subject are not accounted for;
4. To obtain, as useful resources are not located as a consequence;
5. To explore, as it is not possible to, e.g., browse around related topics such as through using related terms in a thesaurus, or see narrower and broader terms or classes, in order to understand the relationships between various nomens for an entity; and, as it is not possible to explore correlations between nomens for the same entity in different controlled vocabularies, e.g., finding a thesaurus descriptor which corresponds to a classification number.

5.1.3 Bibliographic services

Established bibliographic objectives to ensure subject access for journal articles are not adequately supported in large bibliographic services such as Scopus, particularly for the humanities (Golub et al. 2020). For example, Scopus does not use any controlled vocabulary for any humanities discipline (e.g., Arts and Architecture Thesaurus) whatsoever, thus preventing effective retrieval. This finding is well in line with East (2007), who established that ten individual databases in the humanities that he studied provided no controlled vocabularies for humanities resources. Only a minor portion of all articles in the study (Golub et al. 2020) have any controlled vocabulary terms assigned in Scopus; those that do use index terms do so by relying on controlled vocabularies from outside the humanities (EMTREE; MeSH, GEOBASE). The findings also demonstrate the lack of mapping between the vocabularies, which produces duplicates and renders the users unable to use terms from one vocabulary across all the resources.

5.1.4 Repositories

Subject access in repositories of academic pre-prints, articles, and related outputs normally does not rely on KOS. For example, the National Library of Sweden (2019) provides guidelines for repositories that are made available in the Swedish national repositories service, SwePub (http://www.swepub.se), which includes the repository studied here. According to the guidelines, the National Subject Category, used mainly for statistical purposes by the Swedish statistics agency (Statistics Sweden 2016), is the obligatory metadata value to choose from, while keywords do not have a pre-defined value set. While author keywords may be complementary, authors are not trained in indexing, nor are they provided with any indexing guidelines (Golub et al. 2020). Training and guidelines should be provided to the authors to enhance and speed up the depositing process.

5.1.5 Archives

A study conducted in Croatia, Finland, and Sweden in 2016 (Faletar et al. 2017) surveyed archives regarding the interoperability of their metadata. The study found that while archives believe interoperability is important for their institutions and useful for their users, the current level of interoperability is low. This is due to a lack of interest in interoperability at the strategic and managerial level. In addition to the obvious problems with insufficient resources and expertise, at least a part of the reason for the low priority of interoperability can be explained by a similar inertia of established institutional practices described by Bourdenet (2012). Lim and Liew (2011) also concluded that in New Zealand archives did not prioritize metadata sharing. It seems that interoperability was not, in practice, a strategic concern for the majority of the respondents. In addition, the respondents that directly referred to interoperability put this down to the low priority and the lack of interest by managers in their institutions, the lack of a common strategic vision and mutual understanding and collaboration, and the lack of uniform procedures and "rules of the game."

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

There were also problems with technology and standardization that could be traced back to the lack of interest at the strategic and managerial levels.

The discrepancy between the theoretical importance and practical neglect of interoperability of metadata can be framed as a political issue of what is considered to be important in the context of archival work both within archival profession (e.g. in the context of the debate on participatory archives, Huvila 2015; Theimer 2011) and in the society at large (Feather 2013). In addition to the priorities of archival work, it also provides a key to understanding how the concept of interoperability functions as a part of archival practice. Following Pickering (1995), it is possible to make a distinction between the lack of conceptual agency (choosing methods, developing meanings and relations between concepts and principles) and a collision of several disciplinary agencies (applying established methods to solve problems) in how the respondents refer to interoperability. Even if the references to interoperability could be seen as a vague instance of conceptual agency of defining the priorities of specific aspects of archival work and choosing methods how to best reach the users of archival holdings, the influence of the disciplinary agency of digital library, knowledge organization, information retrieval and Semantic Web research (i.e. using the established methods of these fields to solve archival problems versus trying to develop a new better, contextually more appropriate approach) is very apparent.

The authors suggest the need to place more emphasis on exercising conceptual agency related to digital interoperable online archives to overcome the currently unsolved contradiction between the established disciplinary agency of archival work and the disciplinary agencies of related but conceptually and intellectually separate disciplines of knowledge organization, digital libraries, Semantic Web, information retrieval and others. A relevant follow-up question is to what degree archival work needs to be configured according to the demands of interoperability. Considering the significance of specific local contexts, specific uses and users, and the underrated and if problematic, often still viable offline access to individual collections, it is evident that the conceptual agency needs to be exercised with care in order to avoid breaking something that works at least in some respects.

In cross-search services, the most common issues affecting subject searching today are the inconsistency and incompleteness of metadata and the blending of controlled vocabularies, free keywords and full-text indexing (Dempsey 2012; Fagan 2011; Golub, 2016). Interoperability has been acknowledged as a key issue in cultural heritage contexts (Koutsomitropoulos et al. 2012; Seadle 2010). A large number of national and international infrastructure projects are working on making cultural heritage collections interoperable with each other. Semantic Web standards and interoperability opportunities for cross-institutional searching and linking of cultural heritage data have been available for some time now, and many institutions today provide metadata and/or digital information objects to portals such as Europeana and World Digital Library that allow cross-searching of dispersed collections.

Europeana is a prominent cross-search service that combines metadata from thousands of libraries, archives, and museums. The objects are described using different metadata standards, languages and indexing policies. To address the problem, Europeana developed a data model EDM (Europeana Data Model) based on 15 elements of the Dublin Core Metadata Standard, enriched with an additional 13 elements (Europeana Foundation 2017, 2021). Three of the metadata elements are subject related: dc:subject [the subject of the Cultural Heritage Object (CHO)], dc:type (the nature or genre of the CHO), and dc:coverage (the temporal and spatial subjects of a resource). This has been identified as insufficient since it leads to inaccurate search results with high recall and low precision (Gaona-García et al. 2017; Dobreva and Chowdhury 2010). More recently, Europeana decided to adapt EDM to Schema.org (Freire et al. 2020) because it is supported by major Internet search engines (Wallis et al. 2017). The model uses semantic description languages using resource description framework (RDF) and simple knowledge organisation system (SKOS) and links resources according to LOD principles (Gaona-García et al. 2017).

Work on standards like CIDOC-CRM (CIDOC Conceptual Reference Model; CIDOC stands for the International Committee for Documentation of the International Council of Museums) and FRBRoo (FRBR object oriented) is meant to enable the sharing of metadata across institutions, with the idea of creating a one-stop shop for all potentially relevant resources. Europeana is perhaps the most comprehensive example of this idea coming to fruition. It is, therefore, especially important that an FRBRoo-EDM application profile is developed (Doerr et al. 2013).

5.1.6 Museum catalogues

Related research has shown that while some online museums support subject access (Liew 2004), Trant (2006) states that others (especially art museums) do not because providing subject access is not necessary for its operation (unlike object registration, inventory, location control, etc.). In fact, the overall impression seems to be that many museums describe their collections in far too simple terms which include the title of a work, the creator's name, dimensions and sometimes a picture of the museum's object (Fortier and Ménard 2018); this in spite of the standards and guidelines outlined above (ISO 1985; Baca et al, 2006).

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

While a general overview of functional requirements for digital museum search interfaces is lacking, in part due to differences between museum types, an example by the National Gallery of Art lists a total of 75 functional requirements (The Getty Foundation 2012), of which 31 requirements are related to search/retrieval. The requirements include field-based keyword search, auto suggestion of available terms, Boolean operators, refinement of search results by modifying the search criteria, preserving search history and allowing combination/modification of earlier search/browse sessions with the option to add and subtract browse/search facets into the current or past browser search result; support for search/browse functionality with a synonym ring, authority files and provision of alternatives to those entered by the searcher; expansion of results with broader terms; faceted browse searching on criteria which include ofness, aboutness, tag clouds, object type, etc.; linked terms from search results to other results linked to the same terms; controlled vocabularies including at least Union List of Artist Names (ULAN), The Art & Architecture Thesaurus (AAT), Thesaurus of Geographic Names (TGN), ICONCLASS; highlighting keywords from the search phrase in the results. Other relevant functionalities include providing contextual help to users (display of a pop-up short description upon hovering over a function) as well as a display of a visual timeline of artists and works of art and other world events; visualization of artists, artworks and world events on maps (GIS – geographic information systems). In addition, a number of functionalities related to display, ranking and navigation of search results are listed.

A study from 2022 (Golub et al.) assessed the websites of 91 museums, all of which were found to provide online access to at least some of the holdings in Swedish museums and 9 cross-search services. The study analyzed the search interfaces against a set of 21 criteria and showed that effective subject access is largely unavailable in existing services. Few of these support hierarchical browsing of classification schemes and other controlled vocabularies with hierarchical structures, few provide end-user-friendly options to choose a more specific concept to increase search precision, suggest a broader concept or related concepts to increase recall, disambiguate homonyms, or find which term is best to name a particular concept.

In fact, not a single confirmed case of an established subject-related controlled vocabulary in these services was found. This also makes cross-searching across combined databases very challenging, since there is no such control within individual databases, let alone any mapping between vocabularies across the databases. There seem to be efforts underway to alleviate this: KulturNav (https://kulturnav.org) is envisioned as a platform for creating, managing, and distributing linked open name authorities and vocabularies for cultural heritage.

There is a strong need for the implementation of established controlled vocabularies in museums more widely, not only in Sweden. The heterogeneity of object types and the uniqueness of museum materials are a factor in the underuse and even underdevelopment of terminology for the techniques, types and functions of these objects and consequently for their subjects. Even the AAT, the most comprehensive thesaurus for the cultural heritage domain, is constantly evolving through the addition of new concepts. The AAT is multilingual, and translation projects into many languages are currently active, so the need to translate concepts and definitions into Swedish should be emphasized in particular here. In addition, it is important to record unique local terminology in ethnographic museums and museums of local communities more widely, which goes beyond the scope of the AAT.

6.0 Subject functionalities of interfaces

In order to alleviate these problems, library catalogs and related information retrieval systems should employ a number of subject functionalities on their information search interfaces. The search interface should make use of controlled subject terms from vocabularies such as subject headings systems, thesauri and classification systems, to help the user to, for example, choose a more specific concept to increase precision, a broader concept or related concepts to increase recall, to disambiguate homonyms, or to find which term is best used to name a concept. This includes suggesting terms from the controlled vocabularies to support dynamic and interactive search query reformulation and expansion as well as exact and partial matches to the user query terms. The interface should also provide an option to browse hierarchical classification schemes and other controlled vocabularies with hierarchical structures, to help the user further their understanding of the information need and provide support to formulate the query more accurately. Multilingual support based on controlled vocabularies to conduct multilingual searching and browsing should also be supported (Shiri 2012). Finally, the systems should provide interactive online help and instruction on information searching, in order to teach users about search strategies, search techniques and query formulation.

Literature has pointed to 18 functionalities common across cultural heritage institutions as well as 3 additional image-related ones that are important for collections with images and information objects other than publications (Golub et al. 2021). The authors of this work have consolidated these to include those related to KOS-based searching, browsing and indexing.

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

6.1 KOS-based searching

1. Searching using KOS concepts, including terms in the form of single or compound words, phrases, pre-coordinated headings, and class captions from classification systems. Although classification systems use notation to represent concepts, a user should not be expected to know and search using these class symbols.
2. Searching by individual facets or concepts from KOS that compose a complex term (e.g., a class). This includes the ability to search final and complete built classmarks or pre-coordinated index terms, but also individual, built-in facets of the classes or the index terms (cf. Gnoli et al. 2024).
3. Searching by any combination of individual concepts and facets (as above).
4. Automatic alignment (translation) of user search terms into KOS terms. If the user writes a synonym not used per se, the system automatically translates it into the preferred term denoting the same concept. The system resorts to the KOS to be able to do that.
5. Disambiguation—offering the user different concepts (e.g., are you looking for a Jaguar as an animal or a Jaguar as a car?). See, e.g., Google—when typing 'jaguar,' it retrieves documents about cars, but on the right, it also offers 'for animal, see...'.
6. Linking any index term found in a metadata record to all other metadata records with the same index term. This allows the user to click on the term in the metadata records and directly retrieve all other metadata records with exactly the same term.
7. Searching by major and minor themes represented by KOS, if supported by the indexing policy[1].

6.2 KOS-based browsing

1. Browsing by concepts from KOS, which is especially useful for those new to the document collection. Hierarchically structured concept schemes, such as hierarchical classification systems or information retrieval thesauri, are most beneficial. At the narrowest hierarchical levels there should be a manageable number of information resources -- perhaps not more than several dozen or so. If there are many more, the structure should be further developed to include more narrower concepts[2].
2. Browsing by facets, aspects, and individual concepts from controlled vocabularies, such as individual terms from subject headings, as well as captions and notations representing individual concepts from synthesized classmarks (e.g., in Universal Decimal Classification).
3. Showing narrower terms and broader terms to the search terms. When the user types a search term, narrower and broader terms should also be shown for them to explore and consider choosing a more specific or a more general term and related terms (Tudhope and Binding 2008). It also helps the user with disambiguation[3].
4. Displaying results in systematic order(s). As the function of classification notation is to control the sequential order of concepts (Gnoli 2018), this should also be exploited to present results of a search in a meaningful way, making their examination and selection easier. General subjects will thus precede the more specific ones. The default order of presentation will follow the order in the KOS. However, faceted compounds can also be reordered by privileging a specific facet or interest to the user. Also, as in indexing base themes should have been expressed at the lead of compound classmarks before particular themes, items where the searched concept is the base theme should be presented before those where it is just a particular theme (Gnoli and Cheti 2013).

6.3 Enhancing KOS-based indexing

Some functionalities reflect the need to complement controlled vocabularies with other ways of information retrieval:

1. Searching by words from various metadata elements and full-text.
2. Combining controlled subject searching with searching by other bibliographic fields.
3. Adding, searching and browsing end-user tags. This allows for end-user perspectives and inclusion of most current terms from the literature.
4. Linking concepts from one KOS to other relevant ones. This calls for mapping across KOSs in order to support searching across different databases, including multilingual searching.

6.4 Image-related functionalities

1. Searching by image-related characteristics (e.g. size, colour, layout, orientation -- portrait/landscape).
2. Searching using content-based image retrieval (CBIR) methods (e.g. query by example image).
3. Searching by features enabled by IIIF (e.g. deep Zoom viewing).

6.5 Other search functionalities

These are other functionalities that are not based on KOS but are important for the user, such as:

1. Autocompletion, to help the user speed up the typing and help them type accurately.

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

2. Autocorrection, to help correct mistypes in order to enhance retrieval.
3. Highlighting the search term(s) in the retrieved results. This helps the user identify the context of their search term(s).
4. Searching by Boolean and proximity operators, stemming, truncation, wildcard.
5. Combining previous search formulations. This supports more complex information needs, such as those of researchers, who may need to combine a number of search terms and syntaxes including Boolean or proximity operators, stemming, truncation, wildcards and similar.
6. Help on subject searching.

While the features are mutually interdependent in the search process, exact importance of each feature and its dependencies should be subject of future research such as in the context of information retrieval end-user studies. How to best implement each of the features in terms of usability and interface design should also be studied in collaboration with human-computer interaction experts. Many of the subject-specific functionalities, like query expansion based on controlled vocabularies, currently seem to be limited to experimental interfaces (e.g., Alani et al. 2000; Tudhope et al. 2006), so more research in operational information systems is needed.

6.7 A case study

Finally, a recent project (Queerlit) pointed to the need to implement the above search functionalities in a search interface for a dedicated LGBTQI fiction database in Sweden (Golub et al. 2023). The interface has implemented the following functionalities, using the Queerlit thesaurus:

1. Browsing by subjects. However, the Queerlit thesaurus is characterised by a long list of top terms and only a few hierarchical levels for some terms. This prevents visualising the thesaurus in a browsing tree that would be possible, for example, with some classification systems such as Dewey Decimal Classification.
2. Searching by controlled subjects.
3. Browsing by individual concepts from pre-coordinated terms.
4. Searching by a combination of controlled subject terms.
5. Searching by major and minor index terms, as the Queerlit indexing policies support this possibility.
6. Autocompleting search word with suggestions from the Queerlit thesaurus once the user starts typing.
7. Auto-suggesting controlled versions of entered search terms.
8. Searching by words from various metadata elements and full text.
9. Combining controlled subject searching with searching by other bibliographic fields.
10. Linking each subject access point to its resources.
11. Help on searching.

The following desirable functionalities were not possible to implement for the time being, mostly because the database is part of the Swedish Union Catalogue, *Libris*, with its own limitations:

1. Presenting and browsing excerpts of concept hierarchies, matching search terms, to support disambiguation and broadening or narrowing search. This is planned to be implemented, but exact ways to achieve this are challenging as the interface should not be overcrowded with extensive additional features (this is one option that would, in particular, benefit from UX testing, which we plan to conduct).
2. Highlighting search terms in retrieved metadata and resources. Libris does not currently support this and is not able to invest in developing it at the time of writing.
3. Suggesting corrections of mistypes, which do not exist at the level of Libris. To address most common mistypes, these were added as invisible, non-preferred terms to allow successful retrieval.
4. Linking subject access points from one controlled vocabulary to corresponding concepts in others. Each Queerlit thesaurus term has exact or near matches with Homosaurus, SAO, Swedish Children's Subject Headings (Svenska Barnämnesord) and Library of Congress Subject Headings (LCSH), which is listed in the metadata record for each Queerlit thesaurus term. This should be explored in the future to determine how to best connect the collections in different databases using the same subject index term.
5. Advanced searching by Boolean and proximity operators, truncation of searches, wildcard searches. Boolean operators AND, OR, NOT are supported but only in free-text searching in Libris. Phrase searching with quotation marks, truncation and wildcard searching is also available. Proximity operators are not supported.
6. Combining previous search formulations. This feature has not been implemented but will be discussed if possible to do so in the future.
7. Adding, browsing and searching end user tags. Libris does not allow this and Queerlit is dedicated to high specificity and high exhaustivity in its indexing policies. Additional user tags from services like LibraryThing or creating an additional database for user-entered terms has not been considered yet.

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Finally, the interface and its different features need to be evaluated with user studies focusing on user experience (UX), and many of the search interface characteristics and functionalities are expected to evolve further. Further work is needed to explore search interfaces that fully use the subject metadata assigned to LGBTQI works to maximize their findability and use. Also, more research is needed related to user experience (UX) of the specific interfaces.

Conclusion

The International Society for Knowledge Organization (ISKO) has set up an international working group to develop a set of metadata guidelines for procurement of library management systems (LMS) (Guidelines for the effective use of metadata in discovery systems, https://www.isko.org/stac/). The purpose of the guidelines is to ensure that an LMS enables library users to get maximum value from accessing subject metadata from KOSs. Although the current scope is focused on academic libraries and discovery systems, many of the guidelines will be applicable to other cultural heritage institutions, such as archives, galleries or museums.

In summary, as we witness developments in digital scholarship, it is important to provide quality subject access to a vast range of heterogeneous information objects in digital services. The general objective of subject indexing should be that it allows the user to find anything and everything in the collection (including cross-search collections) that is relevant to a certain topic, and this requires the use of controlled vocabularies to ensure high precision and recall.

With an increased reliance on automation, much of the metadata in library information systems remain unused in search interfaces. Most information systems do not take advantage of controlled vocabularies to support the user with too many or too few resources, as well as with term disambiguation, among other challenges of online searching. The problem is further exacerbated in integrated databases such as discovery systems where inconsistent and incomplete metadata and blending of controlled vocabularies, free keywords, and full-text automatic indexing create significant problems for subject searching (Dempsey 2012; Fagan, 2011). This situation is likely related to the decreasing coverage of KOS in the education of information professionals (Hjørland 2022).

In addition to the action called for above, future research should include digital services for primary sources should also be studied in order to establish the current status and to pinpoint the needs for improvement. Interfaces should be designed and tested to support query expansion, word-sense disambiguation, etc. as discussed in this document, based on specific user needs. All these should include user studies, analyses of real search sessions of humanities scholars and interdisciplinary scholars, as well as university students, cultural heritage professionals and the general public.

This work points to problems that have since long been addressed in the design of controlled vocabularies but are rarely applied in user interfaces of information search systems. It provides guidelines for the design of relevant discovery systems which should make use of the intellectual effort and resources invested into creating controlled subject index terms and indexing languages. Subject access in online information retrieval systems should involve most of the above functionalities. We believe these features need to be common on search interfaces across libraries, archives and museums.

Still, this is not to claim that KOSs are perfect per se; it is well known that they may be slow to accommodate new terms or that they may have structures reflecting a less ideal image of the world – for example, organization of top classes in widely used classification systems like Library of Congress Classification or Dewey Decimal Classification. Complementary approaches such as phenomenon-based classification (see, e.g., Gnoli, 2016), social tagging and automatic indexing thus have their place in subject-based information organization. This work, however, focuses on KOSs because their value has been barely utilized at the level of the search interface, reducing the intellectual effort put into creating the KOS and effort invested into subject indexing merely to string matching (between search terms and terms in metadata records). Mechanisms for meaningful ordering, word term disambiguation, more specific terms, and broader terms for identifying which term denotes a concept throughout the database exist in KOS, and putting them to use is long overdue.

Further research

Extensive further research is needed to address the following topics:

1. How to best implement the different search functionalities at the level of the interface?
2. How can these guidelines be uptaken by vendors and decision-makers in libraries who acquire the software?
3. How could KOSs evolve further in order to better address users' needs with interfaces that make use of the KOS beyond pure string matching? (cf. expressive notations, provision of cross references, interdisciplinary systems...)
4. How do language models such as ChatGPT affect the use of, and benefit from subject metadata for searching?

In addition, options that may also help alleviate issues of subject access include social tagging and automated subject indexing. Further research is needed to determine the level

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

to which it is possible to apply automated subject indexing in the library contexts, as well as to determine the value of those automatically assigned index terms, in combination and comparison with end-user assigned index terms as well as catalogers' assigned index terms in the process of information retrieval by end users. All these and the recommended functionalities for subject access, need to be studied in the context of actual end-user search behavior when it comes to their interaction with relevant information systems.

Endnotes

1. See, for example, the Queerlit interface where a distinction is made between main themes versus minor themes (https://queerlit.dh.gu.se).
2. See, for example, http://scigator.unipv.it/ or https://deweysearchsv.pansoft.de/webdeweysearch/
3. See, for example, Figure 2 at http://www.ukoln.ac.uk/projects/enhanced-tagging/demonstrators/; for query expansion, see also EBSCO's Advanced Searching with CINAH database.

References

Alani, Harith, Christopher Jones, and Douglas Tudhope. 2000. "Associative and Spatial Relationships in Thesaurus-Based Retrieval BT – Research and Advanced Technology for Digital Libraries". In *Research and Advanced Technology for Digital Libraries. ECDL 2000*, edited by José Borbinha, and Thomas Baker, 45-58. Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-45268-0_5

Baca, Murtha, Patricia Harping, Elisa Lanzi, Linda McRae, and Ann Whiteside. 2006. *Cataloging Cultural Objects: A Guide to Describe Cultural Works and Their Images*. Chigaco: American Library Association.

Baca, Murtha. 2004. "Fear of Authority? Authority Control and Thesaurus Building for Art and Material Culture Information". *Cataloging & Classification Quarterly* 38, nos. 3/4: 143-51. https://doi.org/10.1300/J104v38n03_13

Bair, Sheila, and Sharon Carlson. 2008. "Where Keywords Fail: Using Metadata to Facilitate Digital Humanities Scholarship". *Journal of Library Metadata* 8, no.3: 249-62. https://doi.org/10.1080/19386380802398503

Barton, Joshua, and Lucas Mak, L. 2012. "Old Hopes, New Possibilities: Next-Generation Catalogues and The Centralization of Access". *Library Trends* 61, no.1: 83-106. https://doi.org/10.1353/lib.2012.0030

Belkin, N. J., R. N Oddy, and H.M. Brooks. 1982. "ASK for Information Retrieval: Part I. Background and Theory". *Journal of Documentation* 38, no.2: 61-71. https://doi.org/10.1108/eb026722

Bourdenet, Phillippe. 2012. "The Catalog Resisting the Web: An Historical Perspective". *Knowledge Organization* 39, no.4: 276-82. doi.org/10.5771/0943-7444-2012-4-276

Broughton, Vanda. 2023. "Facet Analysis: The Evolution of an Idea". *Cataloging & Classification Quarterly*, 1-28. https://doi.org/10.1080/01639374.2023.2196291

Bundza, Maira. 2014. "The Choice Is Yours! Researchers Assign Subject Metadata to Their Own Materials in Institutional Repositories". *Cataloging & Classification Quarterly*, 52, no.1: 110-18. https://doi.org/10.1080/01639374.2013.852439

Casson, Emanuela, Andrea Fabbrizzi, and Aida Slavic. 2011. "Subject Search in Italian OPACs: an Opportunity in Waiting?" In *Subject Access: Preparing for the Future. IFLA Series on Bibliographic Control vol.* 42, 37-50, *edited by* Patrice Landry, Leda Bultrini, Ed O'Neill, and Sandra K. Roe. Berlin: De Gruyter Saur.

Chickering, F. William, and Sharon Q. Yang. 2014. "Evaluation and Comparison of Discovery Tools: An Update". *Information Technology and Libraries* 33, no.2: 5-30. https://doi.org/10.6017/ital.v33i2.3471

Cossham, Amanda F. 2013. "Bibliographic Records in an Online Environment". *Information Research* 18, no.3: 1-17.

Cutter, Charles Ammi. 1897. *Rules for a Printed Dictionary Catalogue*. Washington DC: United States, Bureau of Education.

de La Tierra, T. 2008. "Latina Lesbian Subject Headings: The Power of Naming". In *Radical Cataloging: Essays at the Front*, edited by K.R. Roberto, 94-102. Jefferson: McFarland & Company.

Dempsey, Lorcan. 2012. "Thirteen Ways of Looking at Libraries, Discovery, and the Catalogue: Scale, Workflow, Attention". In *Catalogue 2.0: The Future of the Library Catalogue,* edited by S. Chambers, 179-202. London: Facet Publishing.

Dobreva, Milena, and Sudatta Chowdhury. 2010. "A User-Centric Evaluation of the Europeana Digital Library BT – The Role of Digital Libraries in a Time of Global Change". In *ICADL 2010: The Role of Digital Libraries in a Time of Global Change* , edited by G. Chowdhury, C. Koo, & J. Hunter, 148-57. Berlin, Heidelberg: Springer Berlin Heidelberg.

Doerr, Martin, Stefan Gradmann, Patrick LeBoeuf, Trond Alberg, Rodolphe Bailly, and Marlies Olensky. 2013. *Final Report on EDM--FRBRoo Application Profile Task Force. Report. Europeana Project*. https://cidoc-crm.org/sites/default/files/TaskfoApplication+Profile+EDM-FRBRoo.pdf

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

East, John W. 2007. "Subject Retrieval from Full-Text Databases in The Humanities". *Portal: Libraries and the Academy* 7, no.2: 227-41. https://doi.org/10.1353/pla.2007.0018

Europeana Foundation .2017. *Europeana Data Model – Mapping Guidelines – v.2.4.*

Europeana Foundation. 2021. "About Europeana". Accessed April 16, 2022. https://www.europeana.eu/en/about-us.

Ellero, Nadine P. 2013. "Integration or Disintegration: Where Is Discovery Headed?" *Journal of Library Metadata* 13, no.4: 311-29. https://doi.org/10.1080/19386389.2013.831277

Emanuel, Jennifer. 2011. "Usability of the VuFind Next-Generation Online Catalog". *Information Technology and Libraries* 30, no.1: 44-52. https://doi.org/10.6017/ital.v30i1.3044

Fagan, Jody Condit. 2011. "Federated Search Is Dead—and Good Riddance!" *Journal of Web Librarianship* 5, no.2*:* 77-79. https://doi.org/10.1080/19322909.2011.573533

Faletar, Sanjica Tanacković, , Koraljka Golub, and Isto Huvila. 2017. "The Meaning of Interoperability and its Implications for Archival Institutions: Challenges And Opportunities in Croatia, Finland, And Sweden." *Iinformation Research* 22(1).

Feather, John. 2013. *The Information Society: A Study of Continuity And Change*. 6th ed. Facet Publishing.

Fortier, Alexandre, and Elaine Ménard. 2018. "What Do Museum Website Users Expect from Linked Open Data?" In *Challenges and Opportunities for Knowledge Organization in the Digital Age. Proceedings of the Fifteenth International ISKO Conference 9-11 July 2018 Porto, Portugal,* edited by Fernanda Ribeiro, and Maria Elisa Cerveira. Baden-Baden: Ergon-Verlag, 900-907.

Freire, Nuno, Glen Robson, John B. Howard, Hugo Manguinhas, and Antoine Isaac. 2020. Cultural Heritage Metadata Aggregation Using Web Technologies: IIIF, Sitemaps and Schema.org. *International Journal on Digital Libraries* 21, no. 1: 19-30. https://doi.org/10.1007/s00799-018-0259-5

Gaona-García, Paulo Alonso, Ana Fermoso-García, and Salvador Sánchez-Alonso. 2017. "Exploring the Relevance of Europeana Digital Resources: Preliminary Ideas On Europeana Metadata Quality". *Revista Interamericana de Bibliotecología* 40, no.1: 59-69. https://doi.org/10.17533/udea.rib.v40n1a06

Gnoli, Claudio. 2014. "Boundaries and Overlaps of Disciplines In Bloch's Methodology Of Historical Knowledge". In *Knowledge Organization in the 21st Century: Between Historical Patterns and Future Prospects*: *Proceedings of the Thirteenth International Isko Conference 19-22 May 2014 Krakow, Poland, edited by Wieslaw Babik.* Advances in Knowledge Organization 14. Baden-Baden: Ergon-Verlag, 129-135

Gnoli, Claudio. 2016. "Classifying Phenomena Part 1: Dimensions". *Knowledge Organization* 43, no. 6: 403-15. https://doi.org/10.5771/0943-7444-2016-6-403

Gnoli, Claudio. 2018. "Notation". In *ISKO Encyclopedia of Knowledge Organization*, edited by Birger Hjørland, and Claudio Gnoli. https://www.isko.org/cyclo/notation

Gnoli, Claudio, and Alberto Cheti. 2013. "Sorting Documents by Base Theme with Synthetic Classification: The Double Query Method". In *Classification & Visualization: Interfaces to Knowledge: Proceedings of the International UDC Seminar 24-25 October 2013 the Hague, the Netherlands,* edited by Aida Slavic, Almila Akdag Salah, and Sylvie Davies. Ergon: Würzburg*,* 225-232.

Gnoli, Claudio, Douglas Tudhope, Patrícia de Almeida, Romana Barbui, Ceri Binding and Véronique Blot, 2024. Simpler search in a complex world: browsing ethnographic videos by freely faceted classmarks, ISKO UK 2023, *CEUR Workshops*, https://ceur-ws.org/Vol-3661/06-SHORT-CGnoli-ISKOUK2023.pdf

"The León Manifesto". 2007. http://www.iskoi.org/ilc/leon.php

Golub, Koraljka. 2011. "Knowledge Organization Systems". *Technical Foundations*. November 23, 2011. https://www.diva-portal.org/smash/get/diva2:793321/FULLTEXT01.pdf

Golub, Koraljka. 2016. "Potential and Challenges of Subject Access in Libraries Today on the Example of Swedish Libraries". *International Information & Library Review* 48, no. 3: 204-210. https://doi.org/10.1080/10572317.2016.1205406

Golub, Koraljka. 2018. "Subject Access in Swedish Discovery Services". *Knowledge Organization* 45, no.4: 297-309. https://doi.org/10.5771/0943-7444-2018-4-297

Golub, Koraljka, Jukka Tyrkkö, Joacim Hansson, and Ida Ahlström. 2020. "Subject Indexing in Humanities: A Comparison Between a Local University Repository and an International Bibliographic Service". *Journal of Documentation* 76, no. 6: 1193-1214. https://doi.org/10.1108/JD-12-2019-0231

Golub, Koraljka, Pawel Michal Ziolkowski, and Goran Zlodi. 2022. "Organizing Subject Access to Cultural Heritage in Swedish Online Museums". *Journal of Documentation* 78, no. 7: 211-47. https://doi.org/10.1108/JD-05-2021-0094

Golub, Koraljka, Jenny Bergenmar, and Siska Humlesjö. 2023. "Searching for Swedish LGBTQI Fiction : The Librarians' Perspective". *Journal of Documentation* 79, no. 7: 261-79. https://doi.org/10.1108/JD-05-2023-0080

Golub, Koraljka., Jenny Bergenmar, and Siska Humelsjö, S. 2022. "Searching For Swedish LGBTQI Fiction: Chal-

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

lenges and Solutions". *Journal of Documentation* 78, no.7: 464-84. https://doi.org/10.1108/JD-06-2022-0138

Heery, R., Lyon, L., Tsinaraki, C., Brody, T., Koch, T. and Doerr, M. 2006. "Digital Repositories: An Evaluation Study on the Development and Implementation of Community Repositories to Support Research (And Learning and Teaching) Report on DELOS2 Network of Excellence on Digital Libraries, Deliverable 5.1.1, available at: http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=19458833E5FA5CC45117776085BD0E87?doi=10.1.1.101.5976&rep=rep1&type=pdf

Hider, Phillip, and Ying-Hsang Liu. 2013. "The Use of RDA Elements in Support of FRBR User Tasks". *Cataloging & Classification Quarterly* 51, no. 8: 857-72. https://doi.org/10.1080/01639374.2013.825827

Hjørland, Birger. 2021. "Information Retrieval and Knowledge Organization: A Perspective from the Philosophy of Science". *Information* 12, no. 3:135. https://doi.org/10.3390/info12030135.

Hjørland, Birger. 2022. "Education in Knowledge Organization". In *ISKO Encyclopedia of Knowledge Organization* edited by Birger Hjørland, and Claudio Gnoli. https://www.isko.org/cyclo/education

Hunter, Rhonda N. 1991. "Successes And Failures of Patrons Searching the Online Catalog at A Large Academic Library: A Transaction Log Analysis". *RQ* 30, no. 3: 395-402. https://www.jstor.org/stable/25828813

Huvila, ISTO. 2015. The Unbearable Lightness of participating? Revisiting the Discourses of "Participation" in Archival Literature. *Journal of Documentation* 71, no.2: 358-86. https://doi.org/10.1108/JD-01-2014-0012

International Organization for Standardization. 1985. *Documentation – Methods for Examining Documents, Determining their Subjects, and Selecting Indexing Terms*. ISO Standard No. 5963:1985 approved 1985; reviewed and confirmed 2020

Koutsomitropoulos, Dimitrius A., Eero Hyvönen, and Theodore S. Papatheodorou. 2012. "Semantic Web and Reasoning for Cultural Heritage and Digital Libraries". *Semantic Web* 3, no. 1: 1.

Kuhagen, Judy. 2016. "RDA Content in Multiple Languages". *JLIS.It Rivista Italiana Di Biblioteconomia, Archivistica e Scienza Dell'informazione* 7, no. 2: 299-306. https://doi.org/10.4403/jlis.it-11705

Lardera, Marco, Claudio Gnoli, Clara Rolandi, and Marcin Trzmielewski. 2017. "Developing SciGator, a DDC-Based Library Browsing Tool". *Knowledge Organization* 44, no.8: 638-643. https://doi.org/10.5771/0943-7444-2017-8-638

Lee, Boram, and EunKyung Chung. 2016. "An Analysis of Web-scale Discovery Services from the Perspective of User's Relevance Judgment". *The Journal of Academic Librarianship* 42, no. 5: 529-34. https://doi.org/https://doi.org/10.1016/j.acalib.2016.06.016

Liew, Chern Li. 2004. "Online Cultural Heritage Exhibitions: A Survey of Information Retrieval Features." *Program Electronic Library and Information Systems* 39, no. 1: 4-24. https://doi.org/10.1108/00330330510578778.

Lim, Shirley, and Chern Li Liew. 2011. Metadata Quality and Interoperability of GLAM Digital Images. *Aslib Proceedings* 63, no. 5: 484-98. https://doi.org/10.1108/00012531111164978

Majors, Rice. 2012. "Comparative User Experiences of Next-Generation Catalogue Interfaces". *Library Trends* 61, no.1: 186-207.

Mazzocchi, Fulvio. 2018. "Knowledge Organization System (KOS)". *Knowledge Organization* 45, no.1: 54-78. Also available in *ISKO Encyclopedia of Knowledge Organization*, eds. Birger Hjørland and Claudio Gnoli, https://www.isko.org/cyclo/kos

Markey, Karen. 2007. Twenty-five Years of End-User Searching, Part 1: Research Findings. *Journal of the American Society for Information Science and Technology* 58, no. 8: 1071-81. https://doi.org/10.1002/asi.20462

Meadow, Kelly, and James Meadow. 2012. "Search Query Quality and Web-Scale Discovery: A Qualitative and Quantitative Analysis". *College & Undergraduate Libraries* 19, nos.2/4: 163-75. https://doi.org/10.1080/10691316.2012.693434

National Information Standards Organization. 2005. *Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. ANSI/NISO

Z39.29-2005. Baltimore, MD: NISO, approved July 25, 2005; reaffirmed May 13, 2010.

Olson, Nasrine. 2010. *Taken for Granted – The Construction of Order in the Process of Library Management System Decision Making*. PhD diss., University of Gothenburg.

Osborne, Hollie M., and Andrew Cox. 2015. "An Investigation into the Perceptions of Academic Librarians and Students Towards Next-Generation OPACS and Their Features". *Program* 49, no.1: 23-45. https://doi.org/10.1108/PROG-10-2013-0055

Panofsky, Erwin. 1939. *Studies in Iconology: Humanistic Themes in The Art Of The Renaissance*. Oxford University Press.

Patel, Manjula, Traugott Koch, Martin Doerr, Chrisa Tsinaraki, Nektarius Gioldasis, Koraljka Golub, and Doug Tudhope. 2005. "Semantic Interoperability in Digital Library Systems". *Report of the WP5: Knowledge Extraction and Semantic Interoperability. DELOS2 Network of Excellence in Digital Libraries*. http://delos-wp5.ukoln.ac.uk/project-outcomes/SI-in-DLs/SI-in-DLs.pdf

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY

Pickering, Andrew. 1995. *The Mangle of Practice: Time, Agency, and Science*. Chicago, IL: University of Chicago Press.

RDA Co-Publishers. 2010. *RDA: Resource description & access.* Chicago, IL: American Library Association.

Riva, Pat, Patrick Le Bouef, and Maja Žumer, Maja. 2017. *IFLA Library Reference Model: A Conceptual Model for Bibliographic Information*. Internacional Federation of Library Associations and Institutions, Netherlands.

Seadle, Michael. 2010. "Archiving in the Networked World: Interoperability". *Library Hi Tech* 28, no. 2: 189-94. https://doi.org/10.1108/07378831011047604

Shatford, Sara. 1986. "Analyzing the Subject of a Picture: A Theoretical Approach". *Cataloging & Classification Quarterly* 6, no. 3: 39-62. https://doi.org/10.1300/J104v06n03_04

Shatford, Sara. 1994. "Some Issues in The Indexing of Images". *Journal of the American Society for Information Science* 45, no.8: 583-88.

Shiri, Ali. 2012. *Powering Search: the Role of Thesauri in New Information Enviornments*. Medford, NJ: Information Today Inc.

Siegfried, Susan, Marcia J. Bates, and Deborah N. Wilde. 1993. "A Profile of End-User Searching Behavior by Humanities Scholars: The Getty Online Searching Project Report No. 2". *Journal of the American Society for Information Science* 44, no. 5: 273-291. https://doi.org/10.1002/(SICI)1097-4571(199306)44:5<273::AID-ASI3>3.0.CO;2-X

Svenonius, Elaine. 1994. "Access to Nonbook Materials: The Limits of Subject Indexing For Visual And Aural Languages". *Journal of the American Society for Information Science* 45, no. 8: 600-606. https://doi.org/10.1002/(SICI)1097-4571(199409)45:8<600::AID-ASI15>3.0.CO;2-6

Szostak, Rick, Claudio Gnoli, and Maria José López-Huertas, eds. 2016. *Interdisciplinary Knowledge Organization*. Switzerland: Springer Cham. https://doi.org/10.1007/978-3-319-30148-8

Tarulli, Laurel. 2016. "Managing Outsourced Metadata in Discovery System". *Managing Metadata in Web-Scale Discovery Systems,* edited by Louise F. Spiteri, 137-64. London: Facet Publishing.

The Getty Foundation. 2012. *Moving Museum Catalogues Online. An interim report from the Getty Foundation*. Los Angeles, CA. https://www.getty.edu/foundation/pdfs/osci_interimreport_2012.pdf

Theimer, Kate. 2011. "What Is the Meaning of Archives 2.0?" *The American Archivist* 74, no. 1: 58-68. https://doi.org/10.17723/aarc.74.1.h7tn4m4027407666

Tibbo, Helen. 1994. "The Epic Struggle: Subject Retrieval from Large Bibliographic Databases". *The American Archivist* 57, no.2: 310-26. https://doi.org/10.17723/aarc.57.2.f0650763x258t4p5

Trant, J. 2006. Exploring the Potential for Social Tagging and Folksonomy in Art Museums: Proof of Concept. *New Review of Hypermedia and Multimedia* 12, no.1: 83-105. https://doi.org/10.1080/13614560600802940

Tudhope, Douglas, and Ceri Binding, C. 2008. "Faceted Thesauri". *Axiomathes* 18, 211-222. https://doi.org/10.1007/s10516-008-9031-6

Tudhope, Douglas, Ceri Binding, Dorothee Blocks, D., and Daniel Cunliffe. 2006. "Query Expansion Via Conceptual Distance in Thesaurus Indexed Collections". *Journal of Documentation* 62, no. 4: 509-33. https://doi.org/10.1108/00220410610673873

Villen-Rueda, Luis, Jose A. Senso, and Félix de Moya-Anegón. 2007. "The Use of OPAC In a Large Academic Library: A Transactional Log Analysis Study of Subject Searching". *The Journal of Academic Librarianship* 33, no. 3: 327-37. https://doi.org/10.1016/j.acalib.2007.01.018

Walker, Jenny. 2015. "The NISO Open Discovery Initiative: Promoting Transparency in Discovery". *Insights: The UKSG Journal* 28, no. 1: 85-90. https://doi.org/10.1629/uksg.186

Wallis, Richard, Antoine Isaac, Valentine Charles, and Hugo Manguinhas. 2017. Recommendations for the Application of Schema.org to Aggregated Cultural Heritage Metadata to Increase Relevance and Visibility to Search Engines: the case of Europeana. *Code4Lib Journal 36, no.* 1. https://journal.code4lib.org/articles/12330

Wells, David. 2020. Online Public Access Catalogues and Library Discovery Systems. In *ISKO Encyclopedia of Knowledge Organization*, edited by Birger Hjørland and Claudio Gnoli. https://www.isko.org/cyclo/opac

Yang, Sharon Q., and Melissa A. Hofmann. 2011. "Next Generation or Current Generation?" *Library Hi Tech* 29, no. 2: 266-300. https://doi.org/10.1108/07378831111138170

Zeng, Marcia Lei, Maja Žumer, and Athena Salaba, eds. 2011. "Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model". *IFLA Series on Bibliographic Control* 43. De Gruyter Saur.

https://doi.org/10.5771/0943-7444-2024-3-169 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY