Le Simplegadi ISSN 1824-5226

Vol. XXIV-No. 26 November 2026 DOI: 10.17456/SIMPLE-275

This work is licensed under a Creative Commons Attribution 3.0

Andrea Cuna

Discovery Information Systems and Information Literacy (II). Advocating for a More Subject-Based Pedagogy (*Part II*)*

Abstract I: In generale, i cataloghi online delle biblioteche universitarie offrono agli studenti un'interfaccia con un'ampia casella di ricerca. Questo design intende fornire un'esperienza di ricerca simile a quella di Google. Tuttavia, a differenza di Google, i "discovery systems" includono anche la navigazione a faccette, che permette di raffinare i risultati della ricerca. Chi non conosce il dominio d'interesse trova problematico questo approccio, perché necessita di una guida concettuale per comprendere meglio lo spazio informativo prima di eseguire una query. La ricerca esplorativa è un processo aperto caratterizzato da attività cognitive e di apprendimento che possono migliorare le competenze di "information literacy".

Abstract II: University online library catalogues typically provide undergraduate students with a user interface organised around a large search box. This design is clearly intended to mimic a Google-like search experience. However, unlike Google Search, discovery systems also include faceted navigation, which allows users to filter search results according to multiple criteria. This approach is problematic for users who are unfamiliar with the subject domain of interest and need conceptual guidance to consolidate their understanding of the information space they are about to explore in advance of a query. Exploratory search is an open-ended process involving cognitive and learning activities that can improve information literacy skills.

Keywords: Discovery Information Systems, Information Literacy, Search User Interfaces, Faceted Search, Thesauri.

> The brain is hungry not for method but for content,
> especially content which contains generalizations
> that are powerful, precise, and explicit
> (Turner 1991: 123).

Introduction

Extending the discussion initiated in Part I, which examined the treatment of topical terms in current-generation online public access catalogues (OPACs) and discovery systems and

* Continuation of *Part I*, Cuna (2025).

proposed moving beyond the limitations of the search box through a more subject-oriented approach, this second part considers how that approach can be operationalised and how it can promote learning and information literacy. Physically browsing the library stacks is an important means by which many users gain an overview of the library's collections and "serendipitously" find the items that best suit their information needs. However, while serendipitous discovery may truly happen among the stacks, its chance of success depends on what is actually on the shelf at the time of one's browsing (Barclay 2010).

Information resources in high demand are likely to be off the shelf; those in low demand may have been stored offsite to bring relief to overcrowded shelves; some information resources may accidentally be in the wrong place or unidentifiable because of missing labels; others may be missing or remain unaccounted for; information resources encompassing multiple subjects can only be physically shelved in one place; those stored on high or low shelves may be easily missed by the user's eye; some storage areas (e.g., special collections) are off-limits to users (Barclay 2010). Given these limitations, why should the actual physical layout of a library be a useful mental model to bring into the OPAC interface? There are two overriding reasons for this: (1) many users have fully internalised the model, which can be activated in a number of browsing situations; and (2) whenever physical browsing of the stacks is restricted or somewhat jeopardised, browsing in the OPAC in the same way as information resources are actually arranged on the shelves becomes critical (Joudrey & Taylor 2017: 560).

The OPAC and the classified stacks are not two separate things; they complement each other: the OPAC remedies the shortcomings of the classified stacks by grouping together under a single uniform heading like subjects that are scattered across different classes via a network of cross-references and links to that heading or other related subjects placed in different classes; in turn, the classified stacks remedy the shortcomings of the OPAC by providing in-depth recognition access to the full content of the information resources (Mann 2015). Research and practical evidence show that the OPAC-classified stack combination is a powerful device, provided that the semantic network designed to guide end users through the imaginary library on the screen is displayed to them intuitively and meaningfully to avoid the experience of feeling lost and not knowing what to do next.

The exploration of an unfamiliar environment is a complex and dynamic activity that can be very demanding in terms of sensorimotor and cognitive processes. Besides browsing, another spontaneous and ubiquitous activity characteristic of animal exploratory behaviour is *categorisation*. The "ability to categorise and identify environmentally meaningful objects and events accurately is fundamentally important for survival" (Ashby & Lee 1993: 377). For example, in keeping with Bates' berry-picking metaphor (2016), foraging berries in the woods can be fun and a terrific opportunity to incorporate healthy nutrients into one's diet, but also extremely dangerous. Distinguishing poisonous varieties from edible ones is one of the first things that berry foragers must learn when engaging in such activity.

However, since edible and poisonous berries share many prominent features, learning to sort them into the proper category can be a challenging task. Categorisation is the ability to group like things and separate unlike things by selectively abstracting some char-

acteristics and ignoring others (Harnad 2017: 44). "*Categorization is intimately tied to learning*" (Harnad 2017: 22, emphasis in original). It requires successfully exploiting one's background knowledge while quickly acquiring new knowledge. Indeed, it could be fatal for berry foragers to sample all the varieties they encounter in the environment, including the poisonous ones, in order to learn which are edible. Even so, miscategorisation does happen. It occurs when things are assigned to a different category than the correct one, which results in incorrect associations being made between the assigned category and the other categories. Without a mechanism in place to compress the sheer mass of information arriving at our perceptual system, our information-processing capacity would be overloaded and fail. The "task of category systems is to provide maximum information with the least cognitive effort" (Rosch 1978: 28).

In "search interfaces, category systems are the main tool for navigating information structures and organizing search results" (Hearst 2009: 20). In general, category-based information structures have a twofold function: (1) to give an overview of the contents of a collection; and (2) to organise search results (Hearst 2009: 74). Category system structures deployed in search interfaces can be either flat, hierarchical, or faceted (Hearst 2009: 21). Flat or linear category systems with distinct, recognisable categories are probably the easiest for users to understand, but they do not scale well to large content sets as some overlap often occurs between at least some of the categories. Hierarchical category systems are useful when there are too many categories to be shown at one level, but the deeper the hierarchy, the more users are likely to become disoriented. Also, retaining semantic consistency among categories is more difficult when the hierarchy needs to be modified. Faceted category systems offer maximum flexibility with large collections as they allow users to search and browse the content according to multiple categories or facets simultaneously and provide the necessary scaffolding for exploration and discovery (Hearst 2009). However, if "the facets do not reflect a user's mental model of the space, or if items are not assigned facet labels appropriately, the interface will suffer some of the same problems as directory structures" (Hearst 2009: 194-195). A literature review conducted by Jody Condit Fagan (2010) found that, when intelligently designed, faceted navigation is, by and large, well understood by users and preferred over other navigation structures, though it is still unclear to what extent it can improve the users' search performance within present-day library catalogues. Faceted navigation has been used in e-commerce sites for quite some time now, but it is only relatively recently that libraries have given it the attention it warrants.

Revamping Old Catalogues

Driven by increased awareness of end users' general dissatisfaction with traditional library catalogues and challenged by the rise of new information marketplace competitors, in 2006 some American academic libraries took action to revamp their OPACs with the faceted search functionality to better help their patrons accomplish their search and browse tasks (Antelman *et al.* 2006; Breeding 2007). Faceted search has gained momentum and acceptance within the library community ever since, and it is now considered an effective means to overcome the limitations inherent in the so-called query-response paradigm that had underpinned the design and construction of earlier OPACs. Combining direct search with

browsing in a seamless fashion is now increasingly recognised as being more powerful and effective than either searching or browsing alone. In the typical real-world scenario, a faceted search performed on an OPAC begins with the user typing one or more keywords into a search box. Given the query, the system returns a set of matching results alongside a list of relevant facets. These facets represent specific item properties or attributes (e.g., author, format, genre, location, language) extracted from Machine-Readable Cataloguing (MARC) records stored in the library database. By displaying search results alongside facets, users can (1) gain a better understanding of search results; (2) explore results from different dimensions; (3) narrow down the result set according to relevant metadata and associated values; and (4) expand or reformulate a query by dynamic term suggestion (Hearst 2009; Tunkelang 2009; White & Roth 2009).

Experimental evidence suggests that faceted search is especially useful to users with fuzzy needs in mind who require support and guidance in moving towards a more focused search (Mayr *et al.* 2016; Gäde *et al.* 2015; Villa 2013). For users unfamiliar with the complexity of Library of Congress Subject Headings (LCSH), subject facets, which contain scattered data mostly extracted from existing headings in MARC 21 field 650, can play a pivotal role in providing effective and meaningful representations of topical structures. The well-known negative attitude of library patrons towards LCSH headings and traditional subject access has spurred a steady stream of research focusing both on how to rethink subject analysis models and on how to best expose subject (meta)data in present-day OPACs (Pre- vs Post-Coordination 2007). Although it is recognised that in some domains browsable lists of subject headings may be a suitable option to systematically collapse large numbers of subdivisions and simultaneously provide users with high-level overviews (Markey Drabenstott & Vizine-Goetz 1994: 309-310; Arsenault *et al.* 2008), a discrete body of literature has grown out of attempts to develop online systems in which LCSH headings are broken up (faceted) to group topics into categories for post-coordinate displays (Anderson & Hofmann 2006; Chan *et al.* 2001; Dean 2004; McGrath 2007; O'Neill & Chan 2003).

In presenting a careful analysis of this approach, Karen McGrath (2007) has called attention to important problems that have long plagued the (de)construction of LCSH headings, suggesting viable solutions for them. However, the notions of facet and facet analysis entertained in much of the literature mentioned above differ substantially from those to be found in classic facet analytical theory, initiated by Ranganathan (1967) and furthered by the UK Classification Research Group (Broughton 2019; Hudon 2019), which has now become firmly established in UK LIS (Library and Information Science) theory and practice (Broughton 2006: 58). This fundamental difference in understanding the two notions signals a major shift away from facet analysis as a formal method for knowledge organisation and representation to a much looser approach in which the Subject (or Topic) facet is treated in exactly the same way as any other descriptive metadata rather than on its own merits. The consequences of this conceptual shift can clearly be seen in present-day discovery systems, where a flat alphabetical list has become the standard way of displaying subjects. On the one hand, the use of a flat format display may have greatly contributed to simplifying the interface design and enhancing usability. On the other hand, it poses serious challenges to novice users who want to perform exploratory tasks as they cannot navigate and exploit the

semantic structure of a subject domain to make sense of it. Indeed, for this to happen, these users first and foremost need to gain an understanding of where things fit in and what they mean in the bigger scheme.

The above considerations may help discovery systems developers tailor the user interface to align with today's novice users' mental models and preferences. Ideally, the user interface should incorporate a browse-and-search approach to provide a browsable, high-level, topical overview of the whole information space covered by the system from the very beginning of the search journey. In this way, users would be provided with the scaffolding necessary to move smoothly across the information space. For example, they might start with an interest in one aspect of a given topic and branch out to others as they discover new information and learn about the topic in ways that foster and support their intellectual skills and information literacy.

Beyond the Basics

Although more time is needed to fully understand the impact that generative artificial intelligence (GenAI) will have on the development of components that may be incorporated into discovery systems, there is evidence that GenAI tools are increasingly being used by younger generations of students to perform knowledge-centric tasks. The question of how effectively these tools can be incorporated into existing information systems is a topic that is currently being investigated from both a purely technical perspective and a system-oriented socio-technical one (Storey *et al.* 2025). As to the forms and effects that these developments may have on future bibliographic information systems, the literature available to date mainly reports advances grounded in human-in-the-loop approaches. The emerging trends and possibilities identified so far do not yet provide a sufficiently comprehensive basis for assessing the long-term implications of these developments for the semantic indexing of content. However, before discussing these concerns, it is useful to consider the current state of information literacy. In academic libraries, information literacy remains something that is conceptually and practically difficult to assess. In this regard, Walsh's words are still relevant today: with so "many definitions for information literacy, it is amazing that anyone is able to apply it to instruction" (2011: 3-4). This conceptual fluidity may be the result of a widespread misunderstanding. According to Johnston and Webber (2006: 120), information literacy is not "just a new term for library user education, research skills, or generic attributes" but a normative concept: "a clearly described discipline of information literacy can offer a powerful intellectual and pedagogical force for coherence and relevance".

To bring about this vision, we need to get information literacy out of the "trap" into which it has stepped. At present, information literacy is stuck between a rock and a hard place: the "rock is the current conceptions of information literacy that represent information literacy as a skill or competency that is confined to information access and use, and associated with tools such as text or technology" (Lloyd 2011: 279). Under this lens, information literacy is simply understood as a skill-based practice that focuses on training students in basic information-seeking methods and technology know-how, a sort of toolbox with which undergraduates can solve their information problems, provided that the cases and examples discussed by their instructors adequately reflect their own informational circumstances and needs.

This brings us to the heart of the matter: information needs analysis. In the learner-centred approach, the primary task of instructors is to create a learning and training environment based on students' core learning needs. This analysis involves gathering insights into the students' perspectives to align educational content and training methods with the context in which information needs arise (Dorner *et al.* 2015). Learners' information needs can generally be conceptualised as a gap between their current state of knowledge and the level of knowledge they need to succeed in their academic pursuits. In the context of information seeking, needs can conceptually be divided into two categories: targeted needs and learning needs. Although they are two extremes of a continuum within which intermediate situations can occur – as Bates' berry-picking model clearly shows (2016) – these two categories are sufficient to introduce another important distinction, that between lookup needs and exploratory needs. Basically, lookup needs arise when learners have a specific and well-defined target in mind. They perform lookup searches to find discrete pieces of information. In a typical lookup scenario, users type a query that represents their information need into the search box, after which the search engine returns all the results that match their query. Then, they browse through the results until they find what they are looking for (Athukorala *et al.* 2016). Lookup needs, however, represent only a fraction of learners' needs, which begs the question: what happens when learners' needs go beyond simple lookup? To address this problem, the first thing we need to consider is query length. According to Dirk Lewandowski

> in some cases, simple queries consisting of only one word may be based on a complex information need. Still, they show only a low degree of specificity and – unless they are about a concrete information need – often leave room for multiple interpretations. In contrast, it can be assumed that adding further words leads to a specification of the query (2023: 71).

Whereas one-word queries submitted to general-purpose web search engines are generic, meaning that they often do not provide enough context to clearly understand the user's intent, the same queries executed on discovery systems take on a slightly different meaning. In this context, queries shorter than three words are considered exploratory queries, that is, queries involving information problems related to general concepts that are open-ended and multi-faceted (Athukorala *et al.* 2016). Coming to search with a broader information need in mind, exploratory searchers "issue search queries that are imprecise at first" and "browse through the search results and iteratively reformulate their queries using new keywords they discover" (Athukorala *et al.* 2016: 2645). One of the drawbacks associated with this scenario is that searchers often feel lost in the information space because of information overload. This happens because these searchers have a subject information need in mind, that is, a need associated with subject content.

Speaking from a knowledge organisation perspective, terms like "broader" or "generic" can only refer to relationships between classes and entities. Thus, what an information system should ideally do by default is to provide exploratory searchers (aka subject searchers) with a navigable, classified representation of the domain relevant to their query, where subject terms are organised according to their semantic relationships.

In this way, exploratory searchers would be provided with the scaffolding necessary to build confidence and efficacy in traversing a domain and eventually move to more focused searches. This browsing functionality is consistent with the tenets of exploratory search, which place strong emphasis on learning and investigation activities (Marchionini 2006: 42; White & Roth 2009: 13). Unfortunately, hierarchical subject browsing is not a standard functionality built into current discovery systems. By default, subject terms are grouped into a Topic facet and displayed in a flat list ordered by decreasing number of occurrences. While this approach may be effective for known subject searches, it falls short of expectations when it comes to exploratory subject searches.

To recapitulate, the lack of topical browsing functionalities that can effectively support exploratory search tasks means that current discovery systems are neglecting to address the information-seeking needs of undergraduate students, many of whom are exploratory users, at least at some point in their interaction with the system. What, then, can be done to address this gap? Generally speaking, a careful analysis of students' information-seeking behaviour and needs is an essential starting point. If the analysis emphasises the presence of exploratory needs (searching for general topics), it would be helpful to begin with an information literacy instruction session on how controlled vocabularies, especially thesauruses, work. A thesaurus can be described as a controlled set of terms derived from natural language arranged in a known order and structured in such a way that equivalence, homographic, hierarchical, and associative relationships among terms are displayed clearly and identified by standardised relationship indicators that are employed reciprocally (National Information Standards Organization 2010: 166).

Besides promoting semantic consistency in how subject content is represented by indexers, it facilitates browsing of a domain's topical hierarchy by providing exploratory users with a conceptual map which shows how concepts and ideas are related to one another (Foskett 1980: 419). In a nutshell, although a thesaurus does not convey factual knowledge like a textbook or an encyclopaedia, it is a valuable resource for expanding one's understanding of a domain's language and usage.

The Role of GenAI and Large Language Models (LLMs): Conclusions and Outlook

Research into the application of GenAI in cataloguing, classification, and indexing remains in its early stages, and significant obstacles must be overcome before fully automated bibliographic processing becomes feasible. According to Alireza Noruzi, the incorporation of GenAI capabilities into current bibliographic systems holds several potential advantages, but "much of the conversation about [GenAI] use in book cataloguing, classification, and subject indexing is speculative" (2024: 3). Noruzi's study provides a useful snapshot of the developments achieved by early 2024 in AI-assisted knowledge organisation and subject indexing: ChatGPT, Copilot, and Gemini can support semi-automated cataloguing tasks, but the resulting subject indexing and classification outputs remain highly inconsistent when compared with established human-assigned records. In this sense, the paper reinforces the idea that AI tools do not provide a robust basis for semantic indexing yet. Variations in output across platforms and the lack of alignment with controlled vocabularies and classifi-

cation schemes limit their reliability, necessitating rigorous professional validation prior to integration into bibliographic processes.

On the one hand, these experimental findings underscore the limitations of the early exploratory use of GenAI through browser-based interfaces, largely due to the absence of dedicated infrastructure at the time. On the other hand, the same findings confirm that, despite ongoing technological advances, human oversight, intervention, and feedback remain essential for managing process complexity and ensuring data quality. Far from being supplanted by automation, human involvement remains a key component of AI-driven processes. Recent developments in both research and practice continue to underscore the enduring relevance of the human-in-the-loop approach, whereby human expertise remains essential for guiding, validating, and refining AI-generated outputs. The examples that follow illustrate the diverse ways in which this approach has been operationalised across a broad range of contexts and applications.

Kwok Leong Tang and Yi Jiang's study on the validation of AI-generated subject terms through the Library of Congress Linked Data Service presents a hybrid workflow in which LLMs generate candidate LCSH terms that are subsequently checked against controlled vocabularies and validated by cataloguers. While this approach may improve efficiency and support metadata enrichment, it also indicates that semantic indexing remains far from being fully automated: AI-generated subject headings still show significant accuracy limitations, especially in relation to specificity, authorised terminology, complex subdivisions, and nuanced subject representation, thus requiring external validation and sustained human oversight (Tang & Jiang 2025). Eric Feikert's recent discussion of using Microsoft Copilot to enhance MARC field 505 (Formatted Contents Note) suggests that generative AI can support limited metadata enrichment tasks. However, current applications remain a long way from achieving reliable semantic indexing, as outputs can be inconsistent, may omit or modify relevant content, and often require iterative prompt refinement and human verification (Feikert 2026).

Eric H. C. Chow's study presents a skill-based agentic pipeline for Library of Congress Subject Indexing. The paper proposes a modular workflow in which an AI agent decomposes subject indexing into sequential stages – conceptual analysis, quantitative filtering, authority validation, and MARC field synthesis – rather than treating it as a single prompt-based task. While the results suggest closer alignment with professional LCSH practice than earlier single-step methods, the study also highlights the continuing limitations of automated semantic indexing. Tested on a small corpus, the pipeline showed weaknesses in subdivision completeness and synonym handling, and its outputs were presented as draft recommendations requiring review and refinement by a subject cataloguer (Chow 2026). Notable applications in this area include OCLC WorldShare Record Manager and OCLC Connexion.

Recent enhancements to OCLC WorldShare Record Manager include AI-assisted metadata enrichment features for cataloguing workflows: the bibliographic record is processed by OCLC artificial intelligence to generate suggestions based on WorldCat data. According to the documentation, the service is designed to support the assignment of call numbers and subject headings, with the cataloguer retaining responsibility for reviewing, refining, and validating each suggestion before final inclusion in the bibliographic record (OCLC Support

2026a). Used for bibliographic cataloguing and WorldCat record management, OCLC Connexion is supported by AI-assisted features provided through dedicated cataloguing toolkits and macros. Unlike the native web integration of OCLC WorldShare Record Manager, OCLC Connexion implements artificial intelligence via these external tools, which provide classification and subject suggestions directly editable by the cataloguer, reflecting its nature as a professional client oriented toward automation (OCLC Support 2026b).

Taken together, these examples point to two key issues: first, that short-term consistency could be improved through feedback from experienced professional indexers; and second, that GenAI systems may benefit from different forms of adaptation to domain-specific bibliographic data, including continued pretraining, fine-tuning, and retrieval-augmented approaches. In particular, continued pretraining could allow models to internalise domain-specific statistical and linguistic patterns from large-scale bibliographic corpora; fine-tuning could adapt general-purpose language models to established indexing conventions and controlled vocabularies, while retrieval-augmented approaches could enable the dynamic integration of external bibliographic knowledge during inference, reducing the need for full parametric memorisation.

The prospect of leveraging domain-specific bibliographic datasets through these complementary strategies appears promising. However, realising this potential requires more than advances in algorithmic development alone. It also depends on enhancing the structure, functionality, and semantic richness of bibliographic data, both in terms of semantic representation and content modelling. These improvements are necessary to support a tighter integration between knowledge representation, data modelling, and AI-based processing.

From this perspective, the originality of the approach appears to lie in the interaction of three interconnected layers:

- A human layer, represented by feedback from professional indexers;
- A model layer, corresponding to the adaptation of GenAI systems; and
- A data layer, involving the restructuring of bibliographic metadata to enhance its semantic richness and machine interpretability.

Adopting a data-layer perspective invites a reconsideration of some long-standing assumptions in subject indexing.

The principle of coextension and related approaches to subject string construction may need to be revisited in light of the requirements of AI-driven information processing and multi-level knowledge representation. Specifically, it may no longer be appropriate to constrain the representation of "aboutness" to a single coextensive string reflecting the subject content of an information resource, as such representations may be insufficient for supporting computational approaches such as LLMs and retrieval-augmented systems. Instead, a more flexible framework for semantic representation could be developed, one that accommodates multiple levels of topical granularity – including both primary and subordinate topics – while supporting structured machine interpretation and interoperability.

To improve the accuracy and efficiency of automated parsing, the construction of subject strings should ideally be grounded in paradigmatic relationships and transparent facets.

Lastly, to provide GenAI tools with a richer and more robust foundation for self-training, bibliographic record fields for textual data should be semantically expressive and robust enough to clearly define domain concepts and relationships.

BIBLIOGRAPHY

Anderson, James D. & Melissa A. Hofmann. 2006. A Fully Faceted Syntax for Library of Congress Subject Headings. *Cataloging and Classification Quarterly*, 43, 1: 28-42.

Antelman, Kristen *et al.* 2006. Toward a Twenty-First Century Catalog. *Information Technology and Libraries*, 25, 3: 128-139.

Arsenault, Clément *et al.* 2008. Tensions in Cataloging: Observations on Standards and Implementation. *Journal of Information Ethics*, 17, 1: 28-42.

Ashby, F. Gregory & W. William Lee. 1993. Perceptual Variability as a Fundamental Axiom of Perceptual Science. *Advances in Psychology*, 99: 369-399.

Athukorala, Kumaripaba *et al.* 2016. Is Exploratory Search Different? A Comparison of Information Search Behavior for Exploratory and Lookup Tasks. *Journal of the Association for Information Science and Technology*, 67, 11: 2635-2651.

Barclay, Donald A. 2010. The Myth of Browsing: Academic Library Space in the Age of Facebook. *American Libraries*, 41, 6, https://americanlibrariesmagazine.org/2010/05/19/the-myth-of-browsing/ (consulted on 30/05/2026).

Bates, Marcia J. 2016. The Design of Browsing and Berrypicking Techniques for the Online Search Interface. *Information Users and Information System Design: Selected Works of Marcia J. Bates, Volume III*. Berkeley (CA): Ketchikan Press, 195-216.

Breeding, Marshall. 2007. Chapter 3: Endeca. *Library Technology Reports*, 43, 4, https://journals.ala.org/index.php/ltr/article/view/4540/5333 (consulted on 30/05/2026).

Broughton, Vanda. 2006. The Need for a Faceted Classification as the Basis of all Methods of Information Retrieval. *Aslib Proceedings*, 58, 1/2: 49-72.

Broughton, Vanda. 2019. *Facet Analysis*. London: Facet Publishing.

Chan, Lois Mai *et al.* 2001. A Faceted Approach to Subject Data in the Dublin Core Metadata Record. *Journal of Internet Cataloging*, 4, 1/2: 35-47.

Chow, Eric H. C. 2026. A Skill-Based AI Agentic Pipeline for Library of Congress Subject Indexing. *arXiv*, https://arxiv.org/pdf/2605.03537 (consulted on 30/05/2026).

Cuna, Andrea. 2025. Discovery Systems and Information Literacy (Part I). Musings on the Current State of Search Interfaces. *Le Simplegadi*, 23, 25: 167-177, DOI 10.17456/SIMPLE-248 (consulted on 30/05/2026).

Dean, Rebecca J. 2004. FAST: Development of Simplified Headings for Metadata. *Cataloging and Classification Quarterly*, 39, 1/2: 331-352.

Dorner, Daniel G. *et al.* 2015. *Information Needs Analysis: Principles and Practice in Information Organizations*. London: Facet Publishing.

Fagan, Jody Condit. 2010. Usability Studies of Faceted Browsing: A Literature Review. *Information Technology and Libraries*, 29, 2: 58-66, https://ital.corejournals.org/index.php/ital/article/view/3144/2758 (consulted on 30/05/2026).

Feikert, Eric. 2026. The Art of Cataloging Meets GenAI: Enhancing Tables of Contents with Microsoft Co-Pilot. *Technical Services in Religion & Theology*, 34, 1: 11-15, https://serials.atla.com/tcb/article/download/4527/5825 (consulted on 30/05/2026).
Foskett, Douglas J. 1980. Thesaurus. *Encyclopedia of Library and Information Science*. New York (NY): Marcel Dekker.
Gäde, Maria *et al*. 2015. *Supporting Complex Search Tasks*. Allan Hanbury *et al.* eds. *Advances in Information Retrieval*. Cham (CH): Springer, 841-844.
Harnad, Stevan. 2017. *To Cognize is to Categorize*. Henri Cohen & Claire Lefebvre eds. *Handbook of Categorization in Cognitive Science*. Amsterdam-Boston (MA): Elsevier Science, 21-54.
Hearst, Marti A. 2009. *Search User Interfaces*. Cambridge (MA): Cambridge University Press.
Hudon, Michèle. 2019. *Facet*. Birger Hjørland & Claudio Gnoli eds. *Encyclopedia of Knowledge Organization*, https://www.isko.org/cyclo/facet#refL (consulted on 30/05/2026).
Johnston, Bill & Sheila Webber. 2006. As We May Think: Information Literacy as a Discipline for the Information Age. *Research Strategies*, 20: 108-121.
Joudrey, Daniel N. & Arlene G. Taylor. 2017 [4th ed.]. *The Organization of Information*. Santa Barbara (CA): Libraries Unlimited.
Lewandowski, Dirk. 2023. *Understanding Search Engines*. Cham (CH): Springer.
Lloyd, Annemaree. 2011. Trapped between a Rock and a Hard Place: What Counts as Information Literacy in the Workplace and How Is It Conceptualized? *Library Trends*, 60, 2: 277-297.
Mann, Thomas. 2015 [4th ed.]. *The Oxford Guide to Library Research*. New York (NY): Oxford University Press.
Marchionini, Gary. 2006. Exploratory Search: From Finding to Understanding. *Communications of the ACM*, 49, 4: 41-46.
Markey Drabenstott, Karen & Diane Vizine-Goetz. 1994. *Using Subject Headings for Online Retrieval: Theory, Practice and Potential*. San Diego (CA): Academic Press.
Mayr, Eva *et al.* 2016. Visualization of Cultural Heritage Data for Casual Users. *IEEE VIS Workshop on Visualization for the Digital Humanities*, 1, https://www.cvast.tuwien.ac.at/sites/default/files/bibcite/439/PubDat_250950.pdf (consulted on 30/05/2026).
McGrath, Karen. 2007. Facet-Based Search and Navigation with LCSH: Problems and Opportunities. *Code4Lib Journal*, 1, https://journal.code4lib.org/articles/23 (consulted on 30/05/2026).
National Information Standards Organization. 2010. *ANSI/NISO Z39.19-2005 (R2010) Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies*. Baltimore (MD): NISO, https://groups.niso.org/apps/group_public/download.php/12591/z39-19-2005r2010.pdf (consulted on 30/05/2026).
Noruzi, Alireza. 2024. The Use of Artificial Intelligence in Knowledge Organization and Subject Indexing. *Informology*, 3, 1: 1-8, https://www.researchgate.net/publication/383515463_The_Use_of_Artificial_Intelligence_in_Knowledge_Organization_and_Subject_Indexing (consulted on 30/05/2026).
O'Neill, Edward T. & Lois Mai Chan. 2003. FAST (Faceted Application of Subject Terminology): A Simplified Vocabulary Based on the Library of Congress Subject Headings. *IFLA Journal*, 29, 4: 336-342.

Ranganathan, S. Ramamrita. 1967 [3rd ed.]. *Prolegomena to Library Classification*. Bombay: Asia Publishing House.
Rosch, Eleanor. 1978. *Principles of Categorization*. Eleanor Rosch & B. Barbara Lloyd eds. *Cognition and Categorization*. Hillsdale (NJ): Lawrence Erlbaum, 27-48.
Storey, Veda C. *et al.* 2025. Generative Artificial Intelligence: Evolving Technology, Growing Societal Impact, and Opportunities for Information Systems Research. *Information Systems Frontiers*, 27: 2081-2102, https://link.springer.com/article/10.1007/s10796-025-10581-7 (consulted on 30/05/2026).
Tang, Kwok Leong & Yi Jiang. 2025. Better Recommendations: Validating AI-generated Subject Terms Through LOC Linked Data Service. *arXiv*, https://doi.org/10.48550/arXiv.2508.00867 (consulted on 30/05/2026).
Tunkelang, Daniel. 2009. *Faceted Search*. San Rafael (CA): Morgan & Claypool.
Turner, Frederick. 1991. *Rebirth of Value. Meditations on Beauty, Ecology, Religion, and Education*. New York (NY): State University of New York Press.
Villa, Robert *et al.* 2013. *Search or Browse? Casual Information Access to a Cultural Heritage Collection*. Max L. Wilson *et al.* eds. *Proceedings of the 3rd European Workshop on Human-Computer Interaction and Information Retrieval*, 19-22, https://ceur-ws.org/Vol-1033/paper10.pdf (consulted on 30/05/2026).
Walsh, John. 2011. *Information Literacy Instruction: Selecting an Effective Model*. Oxford (UK): Chandos.
White, Ryen W. & Resa A. Roth. 2009. *Exploratory Search: Beyond the Query-Response Paradigm*. San Rafael (CA): Morgan & Claypool.

WEBLIOGRAPHY

OCLC Support. 2026a. *View AI Metadata Enrichment Suggestions*, https://help.oclc.org/Metadata_Services/WorldShare_Record_Manager/Bibliographic_records/MARC_21_view/Use_AI_Metadata_Enrichment_suggestions (consulted on 30/05/2026).
OCLC Support. 2026b. *OCLC AI Enrichment Macro*. https://help.oclc.org/Metadata_Services/Connexion/Connexion_client/Connexion_client_toolkits/OCLC_Cataloging_AI_macro (consulted on 30/05/2026).
Pre- vs Post-Coordination 2007. Pre- vs. Post-Coordination and Related Issues, https://www.loc.gov/catdir/cpso/pre_vs_post.pdf (consulted on 30/05/2026).

Andrea Cuna is Assistant Professor of Archival Science, Bibliography and Library Science at the Department of Languages and Literatures, Communication, Education and Society of the University of Udine. His main research interests include exploratory and faceted search, cultural heritage information systems and digital libraries, Generative AI in information systems and AI-based translation agents.
andrea.cuna@uniud.it