



Information Research – Vol. 31 No. 2 (2026)

Workplace user engagement with a semantic AI-based search tool

Marianne Lykke, Ann Bygholm, and Tanja Svarre

DOI: <https://doi.org/10.47989/ir31263125>

Abstract

Introduction. This study investigates non-professional searchers' engagement with and challenges in using an AI-based workplace search tool. The tool included twelve AI-assigned, domain-specific metadata types and information from internal and external sources. Drawing on AI literacy perspectives, the study identifies competencies and support needed for effective workplace use.

Method. A case study was conducted using an exploratory sequential mixed-methods design. Data included an interview with a corporate information specialist responsible for AI metadata development, a questionnaire with seventy searchers, and semi-structured interviews with nine searchers.

Analysis. Preparatory interviews provided context for analysis. Questionnaire data were examined using uni- and bivariate statistics, while interview data were thematically analysed to explore searcher engagement and challenges.

Results. The metadata types were found relevant for both exploratory and known-item searches. Users generally found the tool easy to use but identified challenges: domain knowledge was required; associated information needed enrichment; AI metadata roles were not immediately clear; and some routines had to be adapted.

Conclusions. Searchers demonstrated AI competencies and were able to manage challenges. Effective use of AI-based search tools requires experience, domain knowledge, an explorative mindset, and continuous interaction. AI competence development can be understood as the coordinated deployment of tangible, intangible, and human resources supporting advanced information-seeking practices.

Introduction

Artificial intelligence (AI) is currently a major topic in both society and research, not least since the dawn of end-user generative AI in 2022. AI has the potential to disrupt human life, e.g., how we work, live, communicate, and get medical help (Elliott, 2019). One task that can be handled by AI is classifying text by means of machine learning, which can take the form of being either supervised (Kadhim, 2019) or unsupervised (Thangaraj & Sivakami, 2018), or a combination of the two (Wu et al., 2025).

Search and metadata assignment are among the scientific areas that have explored the potential of AI for various text-based tasks (Corrado, 2021; Golub, 2021; Jha, 2023; Mwantimwa & Msoffe, 2025; Vrindha & Syamili, 2025), and within different contexts, such as libraries (Asula et al., 2021), scientific databases (Moulaison-Sandy et al., 2021), academic papers (Yang et al., 2023), university repositories (Lowe et al., 2021), datasets (Chen et al., 2020), and tweets (Samuel et al., 2020).

However, little is known about how AI can assist in text-based enterprise search. Contrary to the examples given above, text-based enterprise information is characterised by being heterogeneous, existing in a plethora of formats, and being stored across a diversity of information systems (Gunadi & Albayrak, 2015; Kruschwitz & Hull, 2017). At the same time, enterprise search differs from information searching in established databases, such as library databases, and on the Internet by having a larger share of people search, shorter queries, and few advanced queries (Domke et al., 2018; Svarre, Lykke, & Bygholm, 2024).

Enterprise search has been associated with various challenges in previous research. Language use is a significant issue. In corporate settings, there are often multiple professional languages involved, and since information for problem-solving is often combined with information from multiple professional groups and corporate information sources, it can be difficult to find a comprehensive set of search terms, due to domain-specific language use, geographical variations, and changes over time (Cleverley & Burnett, 2015; Furnas et al., 1987; Nielsen, 2005; Schuff et al., 2016). Metadata quality is another frequent problem. Stocker et al. (2015) identified barriers related to adequacy and completeness of metadata. Other challenges are related to the information searcher. Freund (2015) identified challenges, such as searchers' awareness of information and sources, the abundance of information available, access control, and the amount of time available to search and use information.

The purpose of this study was to get insight into how enterprise searchers used and assessed the support provided by a search tool that is based on AI-generated metadata and contextual information from a diverse set of corporate sources linked to the metadata. The study took place in an international biotechnology company that conducts research, development, and production of industrial biotechnological products. Through several user studies, a need was identified for a tool to support searchers with domain-specific, contextual metadata that could guide searchers to relevant information and extend the search across time and academic domains (Lykke, Bygholm, Søndergaard, & Byström, 2022; Svarre, Lykke, & Bygholm, 2024). Using Scibite, an ontology-based text-mining search engine, a set of twelve company-specific, as well as general biotechnological metadata types, was added to the display of retrieved documents in the form of clickable metadata linking to internal and external information sources. The study examined the functionality of (a) the AI-generated metadata and (b) the information that the metadata linked to in internal and external resources.

Specifically, the aim was to get insight into

- how corporate searchers understood, used, and valued the AI-generated metadata,
- how they judged the usefulness and transparency of the information that searchers were directed to when they clicked on one of the twelve AI metadata types,

- and what literacy challenges they met in the AI-based search.

Related work

Here, we first outline the importance of metadata in enterprise search to elaborate on the specific role of metadata in this particular context. Next, we present AI-based solutions for handling metadata to provide a framework for the case presented later. We focus on AI solutions based on machine learning, since LLM-based solutions represent a very different approach in working with metadata. Lastly, we provide studies that have presented evaluations or user studies to set the stage for the potential of AI-based metadata from an end-user perspective.

Enterprise search and metadata

Enterprise search differs from general Web search, because of the differing characteristics of internal platforms (Fagin et al., 2003; Schymik et al., 2015). This often leads to reduced satisfaction among users, since they are accustomed to rapid improvements in Web search (Townsend & Mathieu, 2018). While many studies have been conducted to understand general search behaviour (e.g., Given et al., 2023), less attention has been given to enterprise search (White, 2020).

Various studies have investigated the characteristics of employees searching for information as part of solving work tasks in their workplace. Departing from different types of internal systems, the literature confirms that different sources are used to locate relevant information for a task. Mathieu (2022) identifies how a diversity of internal and local sources are used on the path towards relevant information for a task.

Several authors point towards the need for contextual information and metadata, when conducting a search. Allard et al. (2009) identified how it is important to engineers that contextual information is available for the information they are looking for, since the context explains and elaborates on the needed information. To illustrate, Freund (2015) showed how document genres could be used to identify documents that would have specific contextual requirements, such as purpose, restrictions, and the like. However, according to Mathieu (2022), metadata initiatives, if existing at all, are often unstructured or ad hoc, and mostly capture limited fields, such as date, version, project affiliation, or authors and/or modifiers.

Cleverley & Burnett (2019) showed how bad or even missing metadata caused dissatisfaction among searchers and emphasised the importance of good titles and metadata for enterprise search. In an earlier paper, Cleverley et al. (2017) identified how unclear metadata can have an impact on the understanding of information quality.

When too many sources are gathered in an internal system, they can cause retrieval challenges for employees. Mathieu (2022) found that many sources cause overlaps between information units, as well as information overload. Furthermore, she identified issues with employees being presented with information in search results that had restricted access for some users, which was also a source of frustration. By contrast, Cleverley & Burnett (2019) identified dissatisfaction, when users were looking for known information that could not be found, because it was not indexed at the time of retrieval. While trustworthiness of internal information is highly valued, the same is the case for easy access (Allard et al., 2009). Thus, restricted access can be one among several factors that lead users to enter the Internet for information that, in some cases, can be easier to access.

AI for enterprise search

As identified in the introduction, AI has been used for a variety of text-based tasks. However, it has also been suggested as a solution for supporting and improving enterprise search. Pointing out the necessity of search expertise to succeed in searching, Cleverley et al. (2017) conclude their paper, by pointing towards machine learning, semantic networks, and tagging as possible

approaches to improving employees' ability to locate information. Later work has made these suggestions more specific. From the literature on the potential in libraries in general, Oyighan et al. (2024) mentions natural language processing (NLP) as a tool for metadata extraction and generation, which can handle fast growing collections, as well as make metadata more responsive to specialised needs.

Kandepu and Harry (2023) focus on content management systems (CMS) in organisations and identify several solutions for supporting users, out of which some are particularly relevant to the current context of enterprise search. One is the use of semantic clustering, as well as different natural language processing (NLP) analyses, for identifying topics and establishing relationships for discovery. Another is using knowledge graphs to identify content relationships on the basis of various data, including metadata and user data.

Evaluations and use of AI in the context of search and metadata

Different studies examine how AI-based solutions are being accepted and used. One line of studies focusses on information professionals and how they accept using AI for tasks that were previously manual. Togia et al. (2026) used a questionnaire to identify differences in use and acceptance among Greek librarians. They reported that only very few librarians used AI for creating and enhancing metadata, and for automated subject indexing. Academic librarians reported to have a higher readiness to incorporate AI in their work, as well as a higher frequency of use, when compared to public librarians. Similar findings were made by Chen and Li (2024), who conducted a survey of metadata professionals' use of AI for cataloguing and metadata creation. The authors found that AI did not have an important role in respondents' work with metadata, and that the majority did not find AI helpful in the quality of their work or in terms of efficiency. On the other hand, Pinar & Cox (2025) used a literature review with fifty-four papers to identify the biggest challenges in using AI in libraries and archives. They found that the challenges related to finances, data management, and ethics were the most frequent, while metadata and indexing occurred less, but was still observed in twenty-five out of fifty-four papers. It appears that AI is still being integrated in work routines, where it has potential, but metadata professionals also see challenges in its use. While libraries and archives are well covered in the literature, no papers have been identified that cover the corporate perspective.

While we have presented the use of AI for various metadata tasks from the metadata professional's perspective, another highly relevant view is to understand the end-user's perspective. How are end-users searching with the help of AI generated metadata? Does it change their search practices? Little focus has been put on this perspective in the literature. No papers report on library or archive end-users. In the enterprise context, two papers were located: one experiment, and one study that included end-users.

Marques & Murphy (2022) conducted an experiment in which they tested how semantic techniques could identify relevant passages for software engineers. Based on variations of word2vec and BERT, the authors used fifty randomly selected Android development tasks and 133 artifacts to perform the experiment using AnswerBot as the baseline. Precision and recall were used as measures of evaluation. The experiment showed that the alternative approaches, and particularly BERT, achieved similar levels of recall as the baseline approach, and thus could hold potential within the software domain.

Zhou et al. (2023) evaluated an inhouse Q&A system in a large tech company, that is based on two BERT models: one for reranking and one for extracting answer spans. The solution was evaluated using logs and interviews with employees. Qualitative evaluation showed that responses could not be assessed from their strict textual expression. Rather, currency, trustworthiness, context, and situation affected the assessment. Further, contingent answers improved the user experience, even though they were not perfect. Thus, the answers contributed with additional

information and alternative access beyond search engine results. The evaluation emphasised the importance of updated datasets for this type of solution, as well as multiple answers, because single answers can lack nuance.

Enterprise search research shows that metadata is crucial for providing context, supporting relevance judgments, and enabling effective information access. AI has been proposed as a way to scale and improve metadata creation and enrichment. Existing studies mainly examine AI use from the perspective of metadata professionals, especially in libraries and archives. These studies report low adoption rates and highlight concerns about quality, ethics, and organisational readiness. The enterprise or corporate context is largely missing from this body of research. End-user studies in enterprise search are rare and mostly focus on AI-supported retrieval or question-answering systems, rather than on metadata itself. Nevertheless, these studies indicate that users value contextual relevance, trustworthiness, and informational cues over strict accuracy. To date, no empirical studies have directly examined how AI-enhanced metadata influences enterprise end-users' search practices, a gap this paper seeks to address.

Conceptual framework

Because AI applications gain traction in many areas, the interest in AI literacy, i.e. what is required to use these applications in a sensible and productive way, also increases. Within recent years, a number of papers and reviews have been published with the aim of identifying what the concept of AI literacy might involve and how best to afford the competencies concerned. In this paper, we drew on different conceptualisations of AI literacy to understand the challenges employees experience when using AI metadata in their search practice. Focus was on AI literacy for non-AI professionals in a workplace setting.

The concept of literacy, traditionally understood as the ability to read and write, also includes numerical literacy, health literacy, financial literacy, media literacy, and digital literacy, to name but a few. An often-used reference to the notion of AI literacy is Long & Magerko (2020), who mention the need to investigate “what new competencies will be necessary in a future in which AI transforms the way we communicate, work, and live with each other and with machines” (p. 1), and they refer to this set of competencies as AI literacy. In their review of the AI literacy of employees at digital workplaces, Cetidamar et.al. (2024) distinguish between studies focusing on AI literacy in education and AI literacy at workplaces, and, furthermore, between AI literacy at the individual level and AI literacy at the organisational level. They note that most studies of AI literacy have been conducted at the individual level and around education. In our attempt to understand employees' challenges and then discuss what can be done to support employees, we need both the individual and the organisational level. The individual level is about the competencies needed for individual members of the workforce to use AI applications in a productive way and the organisational level is about what it takes for an organisation to benefit from applying AI in their business.

An example of a conceptualisation of AI literacy at the individual level is Ng et al. (2021) which, based on a literature review and inspired by Bloom's taxonomy, proposed four types of competencies to characterise different kinds of capabilities required in interaction with AI applications. Bloom's taxonomy (Bloom, 1956) classified cognitive skills into six hierarchical levels, with each level building upon the previous one, and is a widely used framework within the educational area to identify learning objectives that target different levels of complexity independently of specific learning domains. The taxonomy has been revised and expanded several times (Krathwohl 2002; Huitt 2011). The first three aspects of the AI literacy concept proposed by Ng et al. (2021) cluster the six levels from Bloom into three, that is 1) know and understand, 2) use and apply, and 3) evaluate and create. With the fourth aspect, ethical issues, they add an ethical dimension to the concept and include issues, such as transparency, accountability, and fairness.

Based on a revised version of the conceptual framework, Ng et al. (2024) also developed a questionnaire emphasising not only cognitive but also behavioural and attitudinal issues to measure student learning outcomes in AI learning programmes. Other authors working within the educational and individual field have also developed questionnaires and instruments to measure AI literacy. What they had in common was that they were based on some core constructs identified through literature review, sometimes supported by various theoretical constructs, expert testimony, and testing with students or citizens. For example, Laupiciler et al. (2023) based their scale for the assessment of non-experts in AI literacy on “technical understanding,” “critical appraisal,” and “practical application.” Wang et al. (2023) used “awareness,” “use,” “evaluation,” and “ethics” as basic concepts in a scale for measuring AI literacy. Pinski and Benlian (2023) aimed to develop a scale for measuring general AI literacy, and, based on concepts drawn from the information science domain, they suggested seven dimensions grouped in three categories, that is AI actor knowledge, AI process knowledge, and AI experience. There were then a number of suggestions as to what should be included in the education of students and citizens in the field of AI and how this knowledge could be measured. Common to these was that knowledge of the technology was not enough; you also need practical experience and skills that allow you to assess the consequences of using the AI application critically.

Moving to the organisational level, the findings may be considered somewhat similar. That is, acquiring AI technology in itself does not necessarily make a difference to an organisation. In spite of the potential that AI technologies offer, organisations are faced with a number of challenges that prevent them from realising the benefits. AI is far from being a plug-and-play technology, as Fountaine et al. (2019) put it. They emphasised that it is equally important to align culture, structure, and ways of working to support AI adoption. Thus, there is a need for understanding what such an alignment might involve, that is how organisations build AI capability. Mikalef & Gupta (2021) aimed at identifying the necessary organisational resources that will enable a company to build their AI capabilities. They drew on resource-based theory, the basic idea being that “the building of resources facilitates the formation of organisational capabilities, which, in turn, drive performance gains” (Mikalef & Gupta, 2021, p. 2). Building on theoretical insight from resource-based theory, as well as studies outlining the challenges related to AI adoption, Mikalef & Gupta (2021) proposed eight resources that allow organisations and companies to obtain AI capabilities. Using a classification scheme from the resources-based theory domain, the eight resources were grouped in three main categories: 1) Tangible resources, including data, technology, and basic resources, 2) Human resources, including technical skills and business skills, and 3) Intangible resources, including fostering inter-departmental coordination, organisational change capacity, and risk proclivity. The authors underscored that although both tangible and human resources are necessary and a prerequisite, the intangible resources are extremely important for obtaining AI capabilities. These are characterised as more difficult to identify and also more unique because they are developed through a mix of history, people, and processes that characterise the specific organisation.

Cetidamar et al. (2024) focused on employees and aimed at developing a model of the competences needed to utilise AI technologies for achieving company goals. The idea was that strengthening employees’ AI literacy would enable them to participate in the design and use of AI technologies. Furthermore, they stated that AI literacy is “...an organisational-level capability where individual capabilities add up to an organisational level strength...” (Cetidamar et al., 2024, p. 810). Thus, they argued that employees’ collective AI literacy builds an organisational capacity. Based on a literature search and a bibliometric analysis, they suggested that AI literacy for employees should consist of four capabilities, that is technology-related, work-related, human-machine related, and learning-related.

To sum up, AI literacy can be considered from an individual as well as an organisational perspective. The individual view focuses on employees' capabilities in using and interacting with AI-based systems. Although there may be overlaps, the corporate view differs in including how organisations build capabilities and company goals, and how employees can be included in the design and use of AI. Together, the two perspectives offer a complete view on human AI capabilities in workplace contexts.

The case study and the AI-based search tool

The study took place in an international biotechnology company that conducts research, development, and production of industrial biotechnological products and biopharmaceutical ingredients. The search tool was developed to support search for internal information available in a corporate research archive. The digital research archive was established in the 1980s and contained internal documentation about the company's products and R&D work, e.g., lab test reports, memos, guidelines, and project reports. In general, the research archive was used to find information about previous R&D work and historical information about, for example, products, projects, techniques, and industry applications. Specifically, it was used by the searchers to find new application areas for specific products; look up previous results for certain test techniques or materials; or experience with the application of a product in a specific industry.

The documents were uploaded to the research archive, and metadata were registered by the authors. The registration was based manually assigned metadata: document type, title, author(s), project number, industry, product name, external relations, and internal source. It was also possible to attach related documents in the form of PDF files. Access restrictions were specified by indicating allowed readers or groups. The metadata document category (for example, author, project number, industry, and product names) were registered using controlled drop-down lists. The research archive was accessible by the enterprise search system, which was based on SharePoint and, in addition to the research archive, provided access to the corporate information sources intranet, people directory, news, quality documentation database, laboratory system, and social media. The research archive could be searched as part of the enterprise search interface, and via an independent search interface. The search interface facilitated command-based keyword searching and filtering by metadata.

The development of the search tool was initiated in 2020, where a local interview study with eight informants showed that the employees from R&D had difficulty both searching for and assessing the relevance of retrieved documents. The metadata registration was optional and, therefore, all metadata categories were not necessarily linked to all documents. The document titles were not always designed appropriately and contained neither relevant search terms nor words that could be used to assess the content and relevance of the document.

The findings of the interview study were supported by a larger study of enterprise searching practices across different work areas and work tasks in the enterprise search system (Lykke, Bygholm, Søndergaard, & Byström, 2022). The main finding of the study was the identification of an explorative, tracing search technique, where searchers, by use of historical and contextual knowledge, interactively tried out different search paths through the enterprise search system and its sources to put together a set of information that jointly provided the needed information. The needed contextual and historical information could, for example, consist of information about previous research projects or laboratory results or about earlier business application areas where an enzyme has been used or tested. The searchers did not always possess the necessary historical or contextual knowledge and often had to take a detour by consulting a colleague to get the needed information to continue the search and find the desired information for the search task (Svarre, Lykke, & Bygholm, 2024). Thus, an important finding was the need for a tool to support searchers with specific contextual and historical information about, for example,

research projects, lab results, and business application areas to support the tracing search strategy.

Based on these results, the company chose to use machine-learning categorisation to add highly contextual metadata to retrieved information, to make it easy for the searchers to orient themselves quickly in document content and information relevance and to guide the searcher through the added metadata to the needed contextual and historical information to extend the search strategy. By using Scibite, an ontology-based text-mining search engine, a set of twelve company-specific, as well as general biotechnological, metadata types were added to the display of retrieved documents in the form of clickable metadata links (see Figure 1).

The twelve metadata categories were selected by information specialists in collaboration with the vendor (Scibite). Their selection was informed not only by their perceived importance for internal search processes, but also by considerations of practical applicability. One information specialist characterised this approach as “*the art of the possible*”. Consequently, certain types of information recognised as valuable for contextual searching, such as personal names, were excluded from the metadata schema. Within the organisation, individuals are identified primarily by their initials, and the inclusion of such identifiers would have generated an excessive number of non-discriminatory results. By contrast, entities, such as product names and enzymes, proved more amenable to systematic organisation and were, therefore, more readily incorporated into the metadata categories.

PhD thesis: A characterization of cellobiohydrolase Cel7A from *Trichoderma reesei*: kinetics and enzymatic tools for glycoengineering

RA No.	2021-05401-02	Show more versions
Authors	XXXXX (XXXXXXXXXXXX XXXXXXXXXXXX)	
Department	0972	
Registration	March 26, 2021	
Doc. Type	Published Information / Manuscript	
Ext. Author	XXXXX XXXXX	
From source	Research Archive	
Language	gb	

Company (4)
Organism (54)
MyMS number (1)

Patents (1)
Product (1)
Sequolia (2)

Comments: Notification mail sent to BAMK

Files:
02-2021-03800-02_Publication_release_s.d...
02-2021-03800_PhD_thesis_BamK_handed_i.p...

Figure 1. Display of retrieved document with assigned AI-based metadata

The aim was, through machine-learning assigned metadata, to highlight information from the retrieved document that was important both in terms of relevance assessment and formulation of search queries. When the searcher clicked on the assigned metadata, different types of either internal or external information were displayed depending on the metadata category. The added

metadata led, for instance, to internal lab data reports and results for lab study identification numbers appearing in the retrieved document, to taxonomic information about organisms mentioned in the retrieved document from the external NCBI National Center for Biotechnology Information database, to project information and application areas from the internal R&D project database when a project name and/or number was mentioned, or to predefined Google searches for companies whose names were found in the retrieved document. When a searcher clicked on the added metadata for, for example, the ELN lab test database, first a list of related tests was displayed. When the searcher clicked on one of the test numbers, information about the test was shown. The type of information shown was different for the twelve metadata categories (see Table 1). Some information came from internal sources, i.e., the lab database or the corporate product database, and from external sources, i.e., well-known nomenclatures or taxonomies, such as the NCBI. The choice of associated sources and information was decided by the corporate information specialists.

AI metadata	Description and related taxonomies
Companies	All companies we work with, based on data from SalesForce
ELN	Matching on ELN numbers
Enzymes	Based on names/synonyms from EC nomenclature
Lab Equipment	Matching on data from Labservice
Luna numbers	Matching references to other Luna records
My MS	Matching references to My MS records
NN Numbers	Matching on NN numbers
Patent number	Matching on patents from main authorities
Products	Matching on product names from Promis
Projects	Matching on project numbers + their stage e.g. DEV or OPT
Sequoia identifiers	Matching on valid Sequoia identifiers
Organism	Based on bacterial and fungal species from NCBI Taxon

Table 1. The twelve machine learning metadata and related taxonomies.

The development of the tool was carried out in collaboration with Scibite (<https://www.scibite.com/>), which also provided several of the taxonomies that formed the basis for the identification and assignment of metadata. In practice, however, it turned out that these external generic taxonomies, developed to represent a professional domain in general, were not sufficiently specific in relation to the company's use of and approach to categories. Therefore, internally developed taxonomies were primarily used, and even these had to be developed to achieve a sufficient quality in assignment of metadata.

Research method

We used an exploratory sequential mixed-methods design (Clark et al., 2021) to study the use of and challenges related to the search tool. First, we reviewed the AI-based enterprise search tool with the purpose of gaining insight into the twelve types of AI metadata, with particular emphasis on the kinds of information they lead to. We also focused on interface design and functionality in interaction with the AI metadata, specifically investigating the visibility, accessibility, understanding, and usability of the metadata types and their associated information. The study was conducted between 15 January 2022 and 31 December 2022.

This review was followed by an interview with the information specialist responsible for the development of the AI metadata tool. The interview was semi-structured and aimed to illuminate the background and purpose of the AI metadata, the development process, and existing user experiences with the tool. In addition, during the development process, the development team held informal conversations with eight stakeholders, consisting of a mix of internal employees working with digitalization, enterprise systems, and R&D tasks.

Next, we collected qualitative data on concrete end-user engagement and experiences with the AI interface, metadata, and associated information. Based on the findings from the qualitative interview study, we then designed a questionnaire study to assess the range and generalisability of the qualitative findings.

The end-user interviews were also semi-structured interviews. They opened with questions about participants' jobs and organisational positions, their work tasks, and their use of the research archive in their work. Participants were then asked to carry out two recent searches, while explaining in detail their search moves and considerations. We were particularly interested in how they used and assessed the information provided by the AI metadata and the related information sources. Questions concerning the interviewees' affiliations, work practices, and search tasks were designed to elucidate the context in which the AI search tool was employed. The subsequent structured questions were informed by established insights into challenges associated with enterprise search, as well as by the AI literacy literature concerning key competencies in the use of AI technologies. Open-ended questions emerged from observations of, and discussions about, the interactions that occurred during the searches undertaken by the interviewees as part of the interview. The interviews concluded with a general discussion of the role and usefulness of the semantic AI tool.

Interviews took place via MS Teams, with screen sharing enabled while participants carried out the two search tasks. Interviewees were located in their offices at the case company, while the three researchers were co-located in a conference room at the university. The interviews were audio-recorded and subsequently transcribed.

Nine people participated in the study (Table 2). The interviewees were sampled on the basis of two criteria. First, they were required to have used the newly developed AI search tool and the twelve AI metadata elements. Second, the sample was constructed to ensure collective representation of the organisational units and search tasks for which the tool was designed. Sampling and data collection were discontinued at the point at which additional interviews no longer yielded novel insights or emergent themes. Overall, participants worked with research and development of enzyme products (six participants), product safety (two participants), and product quality (one participant). It is important to note that the interviewees possessed limited experience with the newly introduced AI tool, since it had been released only one month prior to the commencement of the study.

Number	Core area	Work area
1	Supply operations	Quality
2	R&D	Formulation
3	Biosolutions	Liquid products
4	R&D	Enzyme activity
5	Supply operations	Product safety
6	Biosolutions	Technical service
7	Supply operations	Product safety
8	Biosolutions	Liquid products
9	R&D	Enzyme activity

Table 2. End-user interviewees.

The transcriptions were read and coded inductively by the three researchers. Initial open coding was conducted individually, after which the researchers discussed and consolidated the identified themes. In addition to identifying themes, the discussion aimed to determine which themes should be investigated further through the questionnaire.

The questionnaire was distributed to 297 research archive users. Participants were selected based on search frequency, with employees who had conducted the highest number of searches during the previous three months invited to respond. The questionnaire consisted of thirty-eight questions, including opinion and assessment questions using ordinal scales, as well as background and factual questions concerning, for example, types of searches conducted in the archive, search frequency, years of employment, and organisational division and function.

Prior to answering the questionnaire, respondents were asked to view a one-minute video introducing the functionality of the AI metadata. This introduction was included, because the research archive search interface displayed two types of metadata. On the left side of the interface, metadata used to filter search results was displayed, while the machine-learning-based metadata was shown on the right side. During the interviews, it became clear that several participants did not distinguish between these two types of metadata. The introductory video was, therefore, included at the beginning of the questionnaire to ensure that respondents understood what was meant by AI metadata.

Seventy users completed the questionnaire in full, corresponding to a response rate of 24%. The questionnaire was prepared and the data collected using the online survey tool SurveyExact. Subsequently, the data were analysed using Excel. Uni- and bivariate statistical analyses were conducted to examine distributions and correlations.

Findings

We start the presentation of findings by describing how the searchers assessed the usefulness of the AI metadata and the associated information in their search activities. Hereafter, we zoom in on user engagement and on what characterised the searchers' interaction with the AI metadata and related information. We present their specific search activities and the challenges that characterised their use of the semantic tool. The study findings are presented in an integrated manner, combining results from the initial qualitative interview study involving nine users and the subsequent quantitative questionnaire study, completed by seventy respondents. Participants in both samples were selected to represent the groups of information searchers for whom the AI tool is intended to provide support in their search.

Usefulness of machine learning metadata and related information

All nine interviewees expressed that the twelve metadata types were relevant in relation to search and that the metadata referred to relevant domain-specific information that they could use for the context-oriented clarification and specification of their search.

I think the metadata can be relevant in many contexts. Here is an example of a document from the research archive that refers to another document through the metadata. Looking at it, I can see that I get information about the relationship between two lab samples that I didn't know were related. So, this leads me to another place that I wouldn't be able to get to through the common metadata filtering. (4, R&D, enzyme activity)

The search constitutes an instructive example of how AI-generated metadata linking data and information across the organisation's information sources over time, facilitated access to both recent and historical knowledge, thereby supporting the understanding and resolution of the specific search problem. The purpose of AI metadata was precisely to enable retrieval across temporal spans, linguistic evolution, and professional domains through the use of extended language models as explained by the interviewee. In the search situation mentioned above, the AI metadata led the searcher to relevant information that gave a deeper understanding of an enzyme from two laboratory experiments conducted at different points in time.

The perceived relevance of metadata was supported by the questionnaire findings, which showed that AI metadata was assessed as useful or very useful (Figure 2). Specifically, metadata representing enzymes at different stages of the lifecycle in different sources were highlighted as valuable within the search process. For example, enzymes under development were represented through identification numbers and LUNA numbers in ELN and LUNA; finished products associated with various business areas were expressed through product and project names in the Product and Project sources; enzymes at a more general level were represented through generic nomenclature within the Enzyme and Organism sources; and information about enzyme quality were contained in the My MS source.

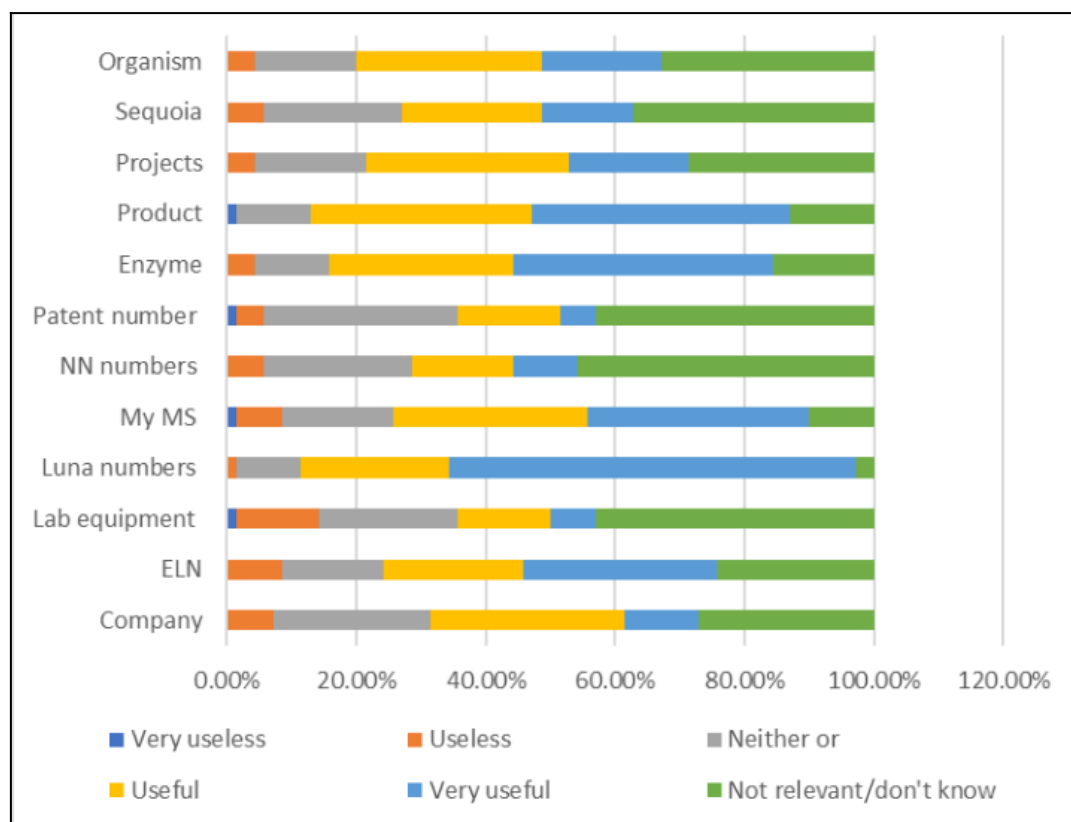


Figure 2. Questionnaire results: Usefulness of AI metadata from the different information sources (n=70).

Questionnaire respondents found AI metadata most relevant in relation to known-item search and specific search (Figure 3). This is likewise reflected in the interviews, in which the interviewees explained that their task is to investigate how a specific enzyme may be applied in other business areas, or how a particular enzyme must be technically adapted and implemented in different contexts. Such search tasks are both highly specific and exhibit characteristics of known-item searching, insofar as the interviewees are aware that prior work concerning the extension of enzyme applications has been conducted and can be retrieved. At the same time, these search tasks are exploratory, since the information seekers do not know precisely which prior work has been undertaken, nor do they possess detailed knowledge of this work in order to retrieve it.

The interviewees further explained that AI-generated metadata was primarily pertinent to research and development-related tasks, where the objective is to explore novel possibilities for an enzyme, rather than, for example, activities within product safety, where the work centres on identifying established findings concerning safety issues associated with a specific product. Within safety-related work, the results were well established and closely tied to the specific enzymes or products under investigation; consequently, such information cannot readily be transferred to other contexts and is, therefore, not considered important to share across tasks in the same manner as in exploratory search activities.

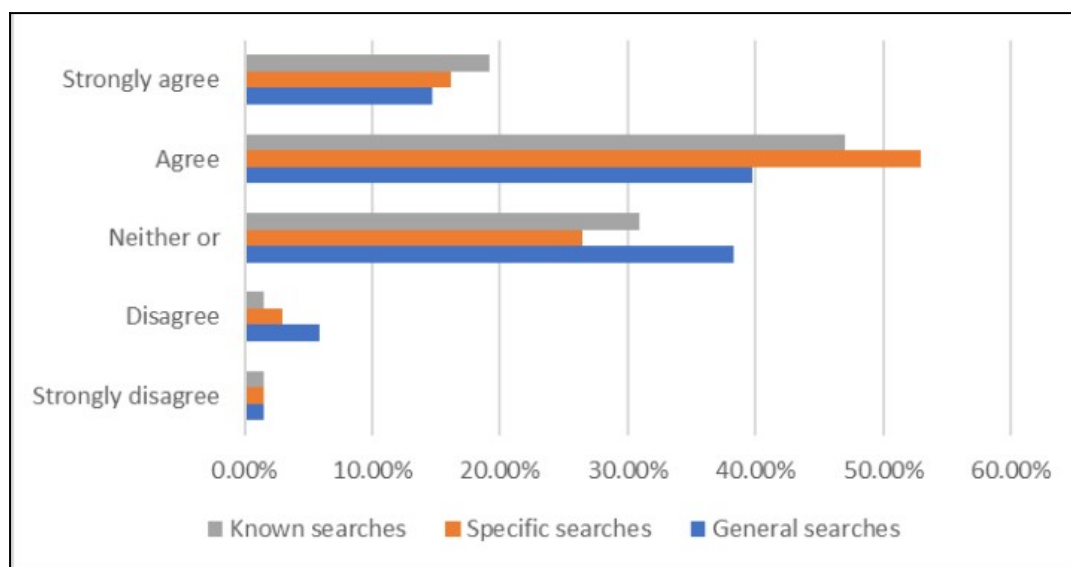


Figure 3. Questionnaire results: Usefulness of AI metadata in relation to search types (N=70)

The interviewees also expressed that the related information was in some cases incorrect in relation to the company's specific perspective and in particular naming of enzymes which, in several instances, diverged from, for example, the nomenclature employed in general taxonomies used by the AI algorithm:

The only enzyme that shows up is an enzyme that is not used as part of the wanted measurement. It is not the enzyme we are interested in. The reason it shows up is because the name that I searched for is an accepted name inside the external enzyme database used for the AI classification, but in the taxonomy, it is accepted for a different enzyme than the one I'm looking for. We use a different name internally. (1, R&D, enzyme activity)

In other search situations, the information was perceived as superfluous and caused more disruption than benefit. In another search demonstration, as part of the interviews, the AI-driven metadata assignment had not accounted for the fact that an enzyme may serve both as the organism under investigation and as an ingredient within a methodological test system. This error recurred across several searches, giving rise to both frustration and diminished trust in the AI tool.

This information is not important in this context because the reagents mentioned are used as part of the method. The reagent is not the substance that we are investigating in this study, just part of the method. (5, supply operations, product safety)

These examples of inappropriate metadata assignment illustrated the importance of ensuring that the taxonomies used for AI metadata reflect the perspective and way of working of the case company. In the first example, the external enzyme taxonomy underlying the AI algorithm classified enzymes based on principles that differed from those used within the company, which resulted in incorrect metadata assignment. In the second example, the taxonomy did not allow for description of the specific role of an enzyme in a study, for example, whether it was an auxiliary substance in the test method or the active substance under investigation. One of the interviewees highlighted that such contextual nuances make it very difficult to automate metadata assignment fully. These examples further demonstrated the importance of information searchers' domain-specific knowledge. Such knowledge is necessary to assess and comprehend both the significance and the limitations of the retrieved metadata and the information to which

they lead. This includes the ability to discern nuances, such as the fact that an enzyme may assume different roles, either as the principal product or as an ingredient within a product.

With regard to practical use and user-system interaction, the questionnaire study showed that respondents agreed that it was clear or very clear how searchers gained access to AI metadata and associated information sources (Figure 4). Respondents also indicated that it was easy to understand that AI metadata provided access to information in other information sources (Figure 5). When comparing the questionnaire responses with the interview data, it becomes evident that the introductory video included in the questionnaire provided a good understanding of the interface functionality. The nine interviewees had not viewed the introductory video, and this lack of introduction was reflected in their assessment of user-system interaction, where they emphasised that the interface was not intuitively understandable and required practice.

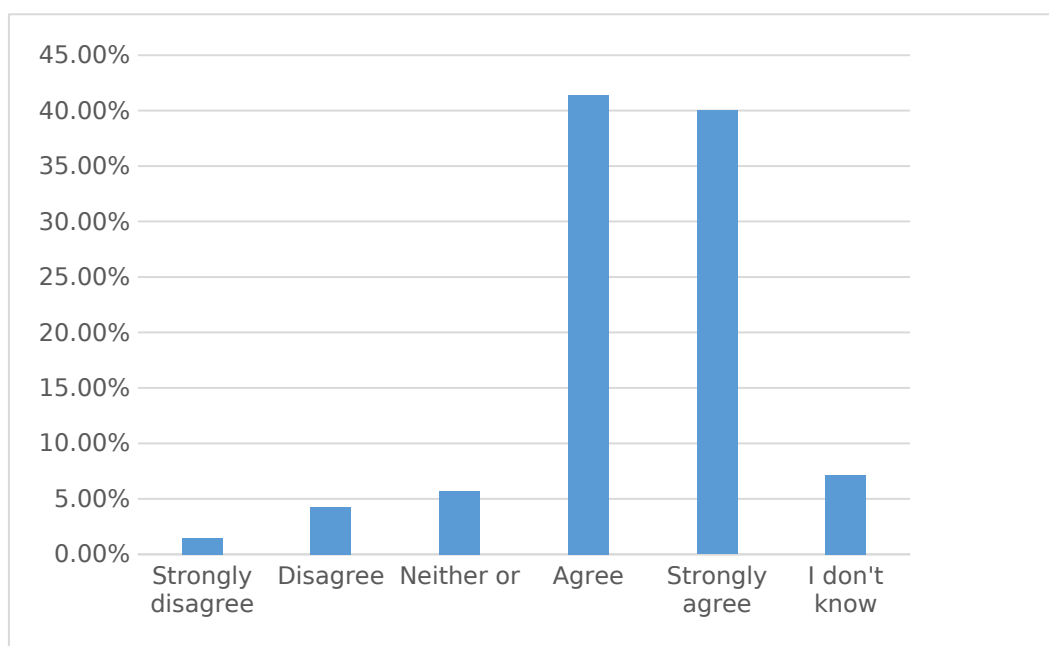


Figure 4. Questionnaire results: Usefulness of access to AI metadata (n=70)

Once the interviewees had tested the new functionality across several searches during the interview, they highlighted direct access to data in other sources as a significant strength:

It exposes me more to this external database, which makes it more likely that I will use it. It increases my awareness that there is this resource to draw on. I think that has a beneficial effect on the research organisation, that these links make it easy to navigate to both internal and external sources. (4, R&D, enzyme activity)

They further emphasised increased visibility of data and efficient, effortless, and fast access:

I think the one with ELN is helpful. In this way you can find your way into the ELN faster. You don't have to open a new browser and search for it yourself. (5, supply operations, product safety).

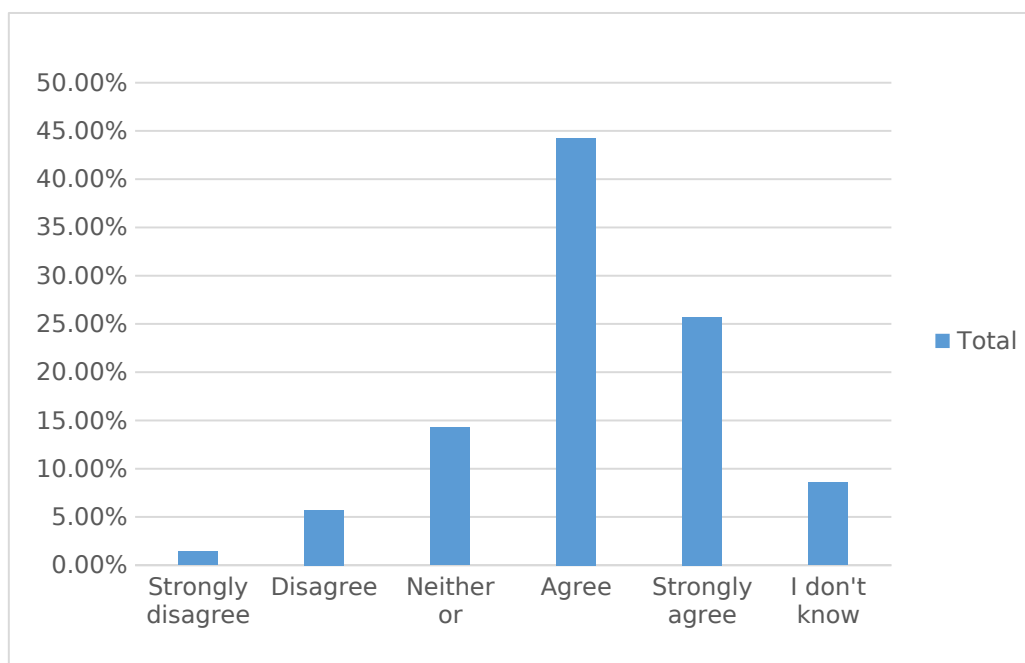


Figure 5. Questionnaire results: Usefulness of access to related information (n=70)

Challenges in engagement with the AI metadata and related information

The interviews provided further insight into the use and usefulness of AI-generated metadata. The interview revealed and provided a more detailed account of two principal challenges that influenced how searchers engaged with and benefited from the AI-generated metadata, namely the significance of domain knowledge and the design of the interface. This section outlines these challenges and examines how the metadata contributed to the search process.

As mentioned above, a good search result required domain knowledge, such as knowing that an enzyme exists in several variations and has been studied across multiple studies over time, in order to take full advantage of the AI tool. Although the purpose of the tool was to provide contextual information to the searcher, understanding the AI metadata and related information still required domain knowledge, as well as knowledge of how to use metadata and information to improve a search. Effective use also required continuous interaction and ongoing critical assessment of search results. Several interviewees expressed that there may be a need to enrich the AI metadata display with explanations of how the metadata could support searching.

Sometimes it seems that this metadata enrichment is pointless, because it requires you to have insight beforehand. I'm just saying that you shouldn't think that you can ever create a system that means that anyone would come in from the street and just sit down and understand what we have been doing for the past many years. It requires some insight. (9, R&D enzyme activity)

The interviewees emphasised that it may be necessary not only to provide explanations for AI metadata, but also to enrich the information presented in the information sources to which searchers are directed. For example, one interviewee explained how the laboratory system only provided a database-specific identification number as a header. This number needed to be supplemented with a title of the laboratory study, including the investigated enzyme, in order to be meaningful to the searcher:

I would have liked a little more information than those rather spartan lists. When I go to the referred information in the lab database, I am not presented with the title. I'm presented with

a database-specific ID that makes no sense to me. We need to get a title, because it is the info that means something to me. (9, R&D enzyme activity)

Another challenge was related to interface design, in which information searchers were presented with filtering via manually assigned metadata on the left side and AI metadata with related information from internal and external sources on the right side. This distinction was unclear to several interviewees. In particular, the difference between the two types of metadata and their respective functions in the search process was not immediately apparent.

There is a difference in functionality where this one links out to an external database and this one does a new search. There is no separation of the different functionalities. They are kind of mixed into each other. The functionality can be helpful, but it is not intuitive until you have used it quite a few times. (6, biosolutions, technical service)

Filtering via metadata was a well-known function, but it was not intuitively understandable that clicking on AI metadata initiated a new search in an internal or external database, thereby providing access to information about related aspects of the original search, such as companies, products, or laboratory results. The interviewees emphasised that this new functionality would require practice over time and the development of new search routines. It also emerged that such misunderstandings initially generated scepticism and reluctance toward the tool.

The AI metadata appeared most relevant for R&D searches, while interviewees from product safety had more difficulty identifying how they could use the functionality. Safety-related searches were described as highly standardised, both because they constitute precise, known-item searches as described above and, from a technical search perspective, because the group of safety employees had developed and inherited practices of manually enriching important documents with contextual information in the title to ensure retrievability. This resulted in safety-related searching being conducted in a systematic and consistent manner. However, during the interview discussions with safety employees, it became apparent that this practice also imposed certain limitations, and that by utilising the new search functionalities and AI-generated metadata, they could extend their use of the enterprise search system. They concluded that the metadata might serve to refine and specify the standardised known-item searches, although this would necessitate the development of new routines. Furthermore, they emphasised that the new search capabilities enabled them to undertake novel types of searches; however, this too would require reconsideration and the establishment of new routines.

There could well be examples where it could be interesting to go in and look at the information linked to AI metadata. For example, the link to these MyMS documents is interesting. A relevant service, but not something that I have missed as such, because in safety we do a recovery batch. We need to be able to document everything and everything needs to be able to be found, so part of the procedure is to add keywords. But I think that it might be a question of getting to know the function. (7, supply operations, product safety)

Discussion

The aims of the study were to gain insight into how workplace searchers used AI-generated metadata and the information associated with it, and to identify literacy challenges in user-system interaction to explore how searchers can be supported. Cetindamar et al. (2024) distinguishes between AI literacy at the individual and organisational level. However, the principal insight derived from the conceptual framework is that experience in the practical use of technology and knowledge is essential for the development of AI capabilities at both the individual and organisational levels. In the discussion, we focus on users' experiences with AI-generated metadata, examining both how such metadata supports and enhances search

processes, how it presents challenges for information searchers, and how they may best be equipped to utilise AI-generated metadata effectively.

Explorative and critical application

The study showed that searchers generally understood and were able to apply the new AI metadata in search. In relation to R&D search tasks, searchers could immediately see how to use AI metadata, whereas searchers from product safety needed to interact with the metadata across different tasks before they could assess its potential.

Searchers working with R&D and application of enzymes had search tasks in which they sought information about a specific enzyme; that is, the task was specific in nature, yet it also had an exploratory dimension, since they did not know precisely what research had been conducted on the enzyme or whether it would be applicable for their purposes. What was crucial for them was to ensure that they identified relevant prior R&D work upon which they could build. They explained how the broad set of AI-generated metadata collectively provided insight into the research conducted over time on a given enzyme, guiding and reminding the information searcher of which sources to consult and which search terms were most worthwhile to pursue further. One interviewee emphasised that the AI-generated metadata functioned as an eye-opener and a shortcut to sources and information.

The searchers from product safety primarily conducted specific, known-item searches and described well-established practices of enriching key documents with tailored titles and metadata to ensure their retrieval. They did not immediately perceive that the new metadata held value in relation to their search tasks. However, upon exploring the AI-generated metadata, they recognised opportunities for broader application but emphasised that this would necessitate the adoption of new search routines.

In the searches, metadata emerged that was incorrect or irrelevant. These errors arose for various reasons. For example, the naming of enzymes varies across companies and between corporate and scientific vocabularies. Likewise, enzymes can assume different roles, such as a fundamental organism or an ingredient. Role distinctions also affected the assignment of company name, where a company might be a partner, supplier, or competitor. These errors primarily caused frustration for the searcher, but they also gave rise to scepticism and doubt regarding the accuracy and reliability of the information. This dissatisfaction and mistrust constitute a well-documented challenge in enterprise search and, regrettably, are not resolved through the application of the AI algorithm in question (Cleverley & Burnett, 2019; Cleverley et al., 2017).

The information searchers emphasised that AI-generated metadata is not a plug-and-play solution. In several searches, they expressed a need for more guidance on how to concretely utilise the metadata and information in their search processes. Fully understanding the metadata and how it could be applied required both domain-specific knowledge and search experience. While AI-generated metadata provides valuable new insights, it does not replace the essential search expertise that an information searcher must possess in enterprise search. It offers a useful new framework for information retrieval, but it does not resolve classical search challenges, such as dealing with diverse and dynamic vocabularies, query formulation, or incorrect metadata assignment (Furnas et al., 1987; Nielsen, 2005; Schuff et al., 2016; Stocker et al., 2015).

The study demonstrated that the participants were experienced searchers with a solid understanding of search algorithms. They approached the AI metadata in an interactive,

explorative and critical manner, relating it to search tasks and prior search experience. They consciously assessed whether the metadata provided new knowledge or support and pointed out relevant problems.

When comparing the interviewees' behaviour with the four AI competencies proposed by Ng et al. (2021, 2024), the participants demonstrated the ability to understand, apply, and evaluate AI metadata by drawing upon their existing search knowledge. In their assessment of AI-generated metadata, they relied on their general understanding of computer-based processes. This was reflected, for example, when they expressed understanding of the AI algorithm's difficulty in distinguishing, for instance, between roles and variations in names. While no interviewees explicitly addressed ethical issues, they implicitly touched on them, by emphasising the importance of domain knowledge and the limitations novices might face when using AI metadata without sufficient contextual understanding.

Similarly, Pinski and Benlian's (2023) distinction of AI actor knowledge, AI process knowledge, and AI experience is reflected in the findings. Although we did not directly assess participants' AI knowledge, the study shows that searchers relied on practical experience to understand and apply AI metadata. They were aware of limitations related to domain-specific nuances that algorithms cannot fully capture. While incorrect metadata assignments were perceived as annoying and sometimes triggered mistrust, the participants' search experience enabled them to manage such issues in practice. It may be argued that they transferred prior experience and existing knowledge of search algorithms to these new AI-based affordances, thereby using prior search experience to develop knowledge of and experience with AI. The fact that the information searchers possessed substantial prior search knowledge meant that they had the capacity to engage in the interaction with AI metadata and the associated information necessary to utilise the new AI-based knowledge to reformulate and improve their searches iteratively. This required the cognitive resources to recognise how different variations in names, as well as knowledge of roles, could be actively employed to identify relevant prior studies concerning a given enzyme.

Overall, the findings indicate that users with extensive search experience and deep domain knowledge were able to use AI-based search systems with the critical approach emphasised by Fountaine et al. (2019). Furthermore, this meant that, on the basis of their existing knowledge, they were able to develop experience with the new AI-based tool, which, in turn, enabled the development of AI literacy.

Ensuring AI use and literacy

Interview participants emphasised the necessity of domain knowledge for their understanding and use of AI-generated metadata, noting that such engagement presupposes prior knowledge. Their reflections centre on the individual user, highlighting that metadata and its associated information may be utilised in markedly different ways, depending on users' personal knowledge and experience.

Extending this line of argument, they further addressed the challenge at an organisational level by proposing functions or tools to support users with domain-specific knowledge. For example, they suggested that non-descriptive identification numbers should be enriched with the full titles and names of the enzymes to which such identifiers refer. In addition, they proposed that the interface should be augmented with explanations and guidance on how the data and information from a given source may be used in search processes. In this sense, the interface ought to be enriched both with contextual knowledge and with search knowledge, indicating how such contextual knowledge may be applied in practice.

Thus, at the organisational level, the study highlights the importance of introducing AI-based systems as tools that require critical use and contextual domain knowledge. Furthermore, the interview participants' suggestions indicate that this knowledge and support should be communicated in multiple forms. Cultures and information-related work practices can be supported through intangible resources, such as inter-departmental coordination, organisational change capacity, and risk proclivity. These resources may enable critical sense-making, the emergence of new information routines, and adaptive practices within the organisation (Mikalef & Gupta, 2021). Continuous and effective interaction with search tools can be further facilitated through human resource initiatives that emphasise the ongoing development of advanced search skills, as well as an increased awareness of the importance of domain knowledge in information searching processes. This may be achieved through introductions and training programmes. The questionnaire's introductory video illustrated how tangible organisational resources (specifically short and easily accessible instructional materials) can support the development of AI literacy and contribute to the adoption and routinization of AI-based systems.

In the interview with the information specialist, they emphasised that the selection of the twelve metadata types was based partly on user studies of search behaviour and on what the company's information specialists, on that basis, assessed to constitute important and relevant contextual knowledge, and partly on what was practically feasible to implement. This entailed, *inter alia*, that person-related metadata linking individuals to, for example, R&D activities, laboratory tests, and projects were not included. None of the interview participants commented on the absence of particular metadata types, nor were any missing metadata types identified in the questionnaire. Since users have not been interactively involved in the development of the AI-based search tool, an important organisational initiative would be to involve users in a future design process, both with regard to the selection of metadata types and the determination of forms of domain and search knowledge to be communicated and incorporated into the interface design, as well as the manner of such communication. For instance, it may be considered whether video-based introductions with illustrative examples are useful, or whether communication should adopt a more dialogic form, for example, drawing on chat-based technologies.

Overall, the study has demonstrated that AI-generated metadata and access to domain-specific information from a broad range of sources are useful. The study further shows that such metadata and associated information transform and, in certain respects, improve the conditions for information retrieval, but that their effective utilisation nonetheless continues to require both domain knowledge and search expertise. There remains a need to develop a range of functionalities and tools that can support information searchers in transferring and developing their prior domain and search knowledge into AI literacy. This development should take place at both the individual and organisational levels.

Study limitations

The present study constitutes a case analysis of an international biotechnology company, drawing upon semi-structured interviews with the corporate information specialist responsible for the development of the AI tool under investigation, a questionnaire study with seventy searchers for whom the tool was developed, and nine semi-structured interviews, including demonstrations of two to three searches conducted using the AI tool by searchers with prior experience of it. The AI tool had only been implemented for one month; consequently, users' experience with the tool was limited. The sampling of interview participants was discontinued at the point at which no further novel insights emerged.

Overall, the study has limited generalisability; however, it provides new and valuable insight into how a group of highly experienced information searchers, possessing substantial domain and

search expertise, interpret and utilise highly specific, contextual metadata and information to expand and reformulate their queries in pursuit of information that is at once precise yet previously unknown. The findings may be considered in relation to similar contexts, in which experienced information searchers are introduced to and interact with new search technologies. Future research should, therefore, seek to examine this particular user group in greater depth and across comparable contexts.

Conclusion

The aim of this study was to investigate how workplace searchers understood and used an AI-based search tool consisting of domain-specific metadata assigned to retrieved documents. Users' assessments and interactions with the tool were examined through interviews with search demonstrations and a questionnaire survey conducted as part of a case study in an international biotechnology company.

The informants found the twelve AI metadata types useful for both exploratory search and specific searches for known information. In terms of user-system interaction, the tool was generally perceived as easy-to-use. However, users also encountered irrelevant and incorrect information, which they found annoying but understandable given the complexity of domain-specific relationships. Several challenges were identified: domain and search knowledge were a prerequisite for effective use; the use of metadata in search was not always readily apparent; there was a need for improved communication of what metadata represent and how they are utilised in search; and new routines were required within established search practices.

During search activities, the interviewees demonstrated the four AI competencies identified by Ng et al. (2021), drawing on existing search experience and understanding of search algorithms, including AI. They were able to understand, apply, and evaluate the AI metadata critically, identify problems, and suggest improvements. The study underscores that users with extensive search experience and deep domain knowledge can use AI-based search systems with the necessary critical approach. Moreover, effective use requires an open-minded, explorative attitude and a willingness to continuously interact with the system to achieve accurate and relevant search results.

The development of AI-related competences can be facilitated through the strategic alignment of tangible, intangible, and human resources, encompassing instructional materials, inter-departmental coordination and organisational change capacity, as well as the ongoing cultivation of advanced search skills and domain-specific knowledge in information-seeking practices.

The insight into how experienced corporate users adopt a novel AI-based search tool—by drawing upon their existing search expertise and substantial domain knowledge, while concurrently extending their overall information-searching literacy to encompass AI-based search technologies—constitutes the principal contribution of the study. The study further provides important insights and practical implications regarding how experienced searchers may be supported, for example, through video-based introductions, explanatory features, and potentially chat-based dialogue integrated within the search interface.

Acknowledgements

The authors would like to thank the company for the generous opportunity to conduct this research and the participants for their engagement and valuable insights on this topic. The authors would also like to thank the anonymous reviewers for their very helpful suggestions in developing the paper further.

About the authors

Marianne Lykke is a Professor at the Department of Culture and Communication, Denmark. She holds a PhD from Åbo Academy in Finland. Her research is focused on user practice studies and user-centred methods to the design and evaluation of information design and knowledge organisation systems. She is particularly interested in information design in workplace environments and museum settings. She can be contacted at mlykke@ikk.aau.dk.

Ann Bygholm is a Professor in Human Centered Informatics at Aalborg University. Her research focuses on how people adopt and adapt IT technologies in their everyday (work) practices, what role these technologies play in people's activities, and how IT can support and enable collaboration, knowledge sharing, and learning in specific situations.

Tanja Svarre is an Associate Professor and co-lead of the research group Purposeful Technology Lab at the Department of Communication and Culture at Aalborg University. Her research interests include professionals' information searching, use, and practice, and evaluation of interactive information retrieval systems. She can be contacted at tanjasj@ikk.aau.dk.

References

- Allard, S., Levine, K. J., & Tenopir, C. (2009). Design engineers and technical professionals at work: Observing information usage in the workplace. *Journal of the American Society for Information Science and Technology*, 60(3), 443–454. <https://doi.org/10.1002/asi.21004>
- Asula, M., Makke, J., Freienthal, L., Kuulmets, H.-A., & Sirel, R. (2021). Kratt: Developing an automatic subject indexing tool for the National Library of Estonia. *Cataloging & Classification Quarterly*, 59(8), 775–793. <https://doi.org/10.1080/01639374.2021.1998283>
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. D. McKay.
- Cetindamar, D., Kitto, K., Wu, M., Zhang, Y., Abedin, B., & Knight, S. (2024). Explicating AI literacy of employees at digital workplaces. *IEEE Transactions on Engineering Management*, 71, 810–823. <https://doi.org/10.1109/TEM.2021.3138503>
- Chen, R.-C., Dewi, C., Huang, S.-W., & Caraka, R. E. (2020). Selecting critical features for data classification based on machine learning methods. *Journal of Big Data*, 7(1), 52. <https://doi.org/10.1186/s40537-020-00327-4>
- Chen, S., & Li, M. (2024). AI for cataloging and metadata creation: Perspectives and future opportunities from cataloging and metadata professionals. *Technical Services Quarterly*, 41(4), 317–332. <https://doi.org/10.1080/07317131.2024.2394919>
- Cleverley, P. H., & Burnett, S. (2015). Retrieving haystacks: A data driven information needs model for faceted search. *Journal of Information Science*, 41(1), 97–113. <https://doi.org/10.1177/0165551514554522>
- Cleverley, P. H., & Burnett, S. (2019). Enterprise search and discovery capability: The factors and generative mechanisms for user satisfaction. *Journal of Information Science*, 45(1), 29–52. <https://doi.org/10.1177/0165551518770969>
- Cleverley, P. H., Burnett, S., & Muir, L. (2017). Exploratory information searching in the enterprise: A study of user satisfaction and task performance. *Journal of the Association for Information Science and Technology*, 68(1), 77–96. <https://doi.org/10.1002/asi.23595>

- Corrado, E. M. (2021). Artificial intelligence: The possibilities for metadata creation. *Technical Services Quarterly*, 38(4), 395–405. <https://doi.org/10.1080/07317131.2021.1973797>
- Domke, E. M., Leidig, J. P., Schymik, G., & Wolffe, G. (2018). Query expansion in enterprise search. *DocEng '18: Proceedings of the ACM Symposium on Document Engineering 2018* (article 33). <https://doi.org/10.1145/3209280.3229111>
- Elliott, A. (2019). *The culture of AI: Everyday life and the digital revolution*. Routledge. <https://doi.org/10.4324/9781315387185>
- Fagin, R., Kumar, R., McCurley, K. S., Novak, J., Sivakumar, D., Tomlin, J. A., & Williamson, D. P. (2003). Searching the workplace Web. *WWW'03: Proceedings of the 12th International Conference on World Wide Web* (pp. 366–375). <https://doi.org/10.1145/775152.775204>
- Freund, L. (2015). Contextualizing the information-seeking behavior of software engineers. *Journal of the Association for Information Science and Technology*, 66(8), 1594–1605. <https://doi.org/10.1002/asi.23278>
- Furnas, G. W., Landauer, T. K., Gomez, L. M., & Dumais, S. T. (1987). The vocabulary problem in human-system communication. *Communications of the ACM*, 30(11), 964–971. <https://doi.org/10.1145/32206.32212>
- Given, L. M., Case, D. O., & Willson, R. (2023). *Looking for information: Examining research on how people engage with information*. Emerald Group Publishing.
- Golub, K. (2021). Automated subject indexing: An overview. *Cataloging & Classification Quarterly*, 59(8), 702–719. <https://doi.org/10.1080/01639374.2021.2012311>
- Gunadi, E., & Albayrak, S. (2015). Information aggregation in an enterprise. In F. Hopfgartner (Ed.), *Smart information systems: Computational intelligence for real-life applications* (pp. 99–122). Springer International Publishing. https://doi.org/10.1007/978-3-319-14178-7_4
- Jha, S. K. (2023). Application of artificial intelligence in libraries and information centers services: Prospects and challenges. *Library Hi Tech News*, 40(7), 1–5. <https://doi.org/10.1108/LHTN-06-2023-0102>
- Kadhim, A. I. (2019). Survey on supervised machine learning techniques for automatic text classification. *Artificial Intelligence Review*, 52(1), 273–292. <https://doi.org/10.1007/s10462-018-09677-1>
- Kandepu, R. K. & Harry, A. (2023). The rise of AI in content management: Reimagining intelligent workflows. *American Journal of Engineering, Mechanics and Architecture*, 1(7), 78–85.
- Kruschwitz, U., & Hull, C. (2017). Searching the enterprise. *Foundations and Trends in Information Retrieval*, 11(1), 1–142. <https://doi.org/10.1561/15000000053>
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *CHI'20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). <https://doi.org/10.1145/3313831.3376727>
- Lowe, D. B., Dollinger, I., Koster, T., & Herbert, B. E. (2021). Text mining for type of research classification. *Cataloging & Classification Quarterly*, 59(8), 815–834. <https://doi.org/10.1080/01639374.2021.1998281>
- Lykke, M., Bygholm, A. K. M., Søndergaard, L. B., & Byström, K. (2022). The role of historical and contextual knowledge in enterprise search. *Journal of Documentation*, 78(5). <https://doi.org/10.1108/JD-08-2021-0170>

- Marques, A., & Murphy, G. C. (2022). Evaluating the use of semantics for identifying task-relevant textual information. *2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)* (pp. 240–251). <https://doi.org/10.1109/SANER53432.2022.00039>
- Mathieu, C. (2022). Defining knowledge workers' creation, description, and storage practices as impact on enterprise content management strategy. *Journal of the Association for Information Science and Technology*, 73(3), 472–484. <https://doi.org/10.1002/asi.24563>
- Moulaison-Sandy, H., Adkins, D., Bossaller, J., & Cho, H. (2021). An automated approach to describing fiction: A methodology to use book reviews to identify affect. *Cataloging & Classification Quarterly*, 59(8), 794–814. <https://doi.org/10.1080/01639374.2021.1992694>
- Mwantomwa, K., & Msoffe, G. (2025). Application of generative artificial intelligence in library operations and service delivery: A scoping review. *Technical Services Quarterly*, 42(2), 139–168. <https://doi.org/10.1080/07317131.2025.2467574>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*, 55(3), 1082–1104. <https://doi.org/10.1111/bjet.13411>
- Nielsen, M. L. (2005). Special libraries and specialized vocabularies in the digital age. In: *Encyclopedia of library and information science*. Dekker.
- Oyighan, D., Ukubeyinje, E. S., David -West, B. T., & Oladokun, B. D. (2024). The role of AI in transforming metadata management: Insights on challenges, opportunities, and emerging trends. *Asian Journal of Information Science and Technology*, 14(2), 20–26. <https://doi.org/10.7012/ajist-2024.14.2.4277>
- Pinar, A., & Cox, A. (2025). An analysis of artificial intelligence (AI) capability in libraries and archives. *Cataloging & Classification Quarterly*, 63(6–7), 566–599. <https://doi.org/10.1080/01639374.2025.2539790>
- Russell-Rose, T., Chamberlain, J., & Azzopardi, L. (2018). Information retrieval in the workplace: A comparison of professional search practices. *Information Processing & Management*, 54(6), 1042–1057. <https://doi.org/10.1016/j.ipm.2018.07.003>
- Samuel, J., Ali, G. G. M. N., Rahman, M. M., Esawi, E., & Samuel, Y. (2020). COVID-19 public sentiment insights and machine learning for tweets classification. *Information*, 11(6), paper314. <https://doi.org/10.3390/info11060314>
- Schuff, D., Louis, R. S., Corral, K., & Schymik, G. (2016). Word ambiguity and search: Implications for enterprise performance management. *The Americas Conference on Information Systems (AMCIS) 2016 Proceedings*. https://web.archive.org/web/20200320101152id_/https://aisel.aisnet.org/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1075&context=amcis2016
- Schymik, G., Corral, K., Schuff, D., & St. Louis, R. (2015). The benefits and costs of using metadata to improve enterprise document search. *Decision Sciences*, 46(6), 1049–1075. <https://doi.org/10.1111/deci.12154>

- Stocker, A., Richter, A., Kaiser, C., & Softic, S. (2015). Exploring barriers of enterprise search implementation: A qualitative user study. *Aslib Journal of Information Management*, 67(5), 470–491. <https://doi.org/10.1108/AJIM-03-2015-0035>
- Svarre, T., Lykke, M. & Bygholm, A. K. M. (2024). Searching for people in the workplace: Aims, behaviour, and challenges. *Information Research: an International Electronic Journal*, 29(2), Article 2. <https://doi.org/10.47989/ir292848>
- Thangaraj, M., & Sivakami, M. (2018). Text classification techniques: A literature review. *Interdisciplinary Journal of Information, Knowledge, and Management*, 13, 117–135. <https://doi.org/10.28945/4066>
- Togia, A., Tsigilis, N., & Koseoglou, E. (2026). Artificial intelligence in Greek libraries: Perceptions, challenges, and readiness for adoption. *Journal of Librarianship and Information Science* online first. <https://doi.org/10.1177/09610006251410240>
- Townsend, R., & Mathieu, C. (2018). Improving enterprise content findability through strategic intervention. *The Code4Lib Journal*, 42.
- Vrindha, K., & Syamili, C. (2025). Navigating the AI landscape in libraries: A PRISMA-based systematic analysis of AI applications in libraries. *Journal of Web Librarianship*, 19(1), 45–61. <https://doi.org/10.1080/19322909.2025.2468697>
- White, M. S. (2020). Information flows – or does it? The complexity of enterprise information management. *Business Information Review*, 37(3), 103–110. <https://doi.org/10.1177/0266382120950114>
- Wu, M., Brandhorst, H., Marinescu, M.-C., Lopez, J., Hlava, M., & Busch, J. (2025). Automated metadata annotation: What is and is not possible with machine learning. *Data Intelligence*, 5(1), 122–138. https://doi.org/10.1162/dint_a_00162
- Yang, H., Wang, N., Yang, L., Liu, W., & Wang, S. (2023). Research on the automatic subject-Indexing method of academic papers based on climate change domain ontology. *Sustainability*, 15(5), 3919. <https://doi.org/10.3390/su15053919>
- Zhou, D. X., Liu, L., Anubhai, A., Shandilya, M., Sigalas, S., Wang, W. Y., & Huang, Z. (2023). Beyond accurate answers: Evaluating open-domain question answering in enterprise search. *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*, 308–312. <https://doi.org/10.1145/3576840.3578314>

Copyright

Authors contributing to *Information Research* agree to publish their articles under a [Creative Commons CC BY-NC 4.0 license](https://creativecommons.org/licenses/by-nc/4.0/), which gives third parties the right to copy and redistribute the material in any medium or format. It also gives third parties the right to remix, transform and build upon the material for any purpose, except commercial, on the condition that clear acknowledgment is given to the author(s) of the work, that a link to the license is provided and that it is made clear if changes have been made to the work. This must be done in a reasonable manner, and must not imply that the licensor endorses the use of the work by third parties. The author(s) retain copyright to the work. You can also read more at: <https://publicera.kb.se/ir/openaccess>