Theorising notions of searching, (re)sources and evaluation in the light of generative AI

Olof Sundin

Abstract

Introduction. The introduction of publicly available large language models (LLMs) since 2022 has significantly challenged traditional library and information studies (LIS) core concepts. This paper argues that LIS needs to rethink aspects of its conceptual framework to address the challenges posed by the proliferation of AI-generated content and how this content is produced.

Analysis. The study employs a theoretical analysis to critically examine the LIS concepts of *search*, *sources* and *evaluation* in the light of an increasingly AI-infused information infrastructure.

Results. The main argument of the paper is that due to changes in the information infrastructure, sources are becoming increasingly invisible for people when they look for information. This has profound implications for how searching for and evaluation of information can be conceptualised.

Conclusion. The paper is concluded by offering conceptual insights on how to theoretically navigate the rapidly evolving information landscape.

Introduction

In his seminal monograph on cognitive authority, the late Patrick Wilson (1983, p. 13) writes, '*All I know of the world beyond the narrow range of my own personal experience is what others have told me*'. If we agree with Wilson, this is why access to and reliance on information sources, and the systems that enable people to find them, are so important. But what will the consequences be for society and individuals if '*what others have told me*' is increasingly generated by AI? I had Wilson's words in mind when I came across an article in the leading Swedish morning newspaper *Dagens Nyheter* in 2023. The article included the following conversation between Swedish pupils in the second year of a Swedish upper-secondary school:

> – *It feels like you get a more direct answer without having to search and scroll. It's like all of Google in one package, says [first pupil].*
>
> – *First you read books, then you googled and in the future you might use GPT to get information, adds [second pupil].* (Wasell, 23 June 2023. Translated from Swedish)

The article was one of many published on ChatGPT at the time. In it, ChatGPT is introduced as both a ghostwriter and a technology for searching, meeting mixed reactions in schools. The quoted pupils were asked to reflect on how they used ChatGPT in their schoolwork. The pupils seem to understand ChatGPT as a replacement for earlier information technologies, leaving the reader of the news article with a chronology of progress: the book, the search engine and ChatGPT. For me, the quote signals that library and information studies (LIS) need to discuss some crucial challenges that arise when AI-generated content becomes more important for people. My primary aim is to contribute to a conceptual discussion of search, sources and evaluation in an information infrastructure increasingly permeated by AI. This is relevant to LIS as a whole, but perhaps especially to subfields such as information behaviour, information retrieval, library studies and – in the particular focus of the paper – information literacy. My main argument is that changes in the information infrastructure make it increasingly difficult for people to recognise sources.

When Brian Campbell Vickery, in his classic 1961 monograph *On retrieval system theory*, distinguishes between finding '*information about documents*' and finding '*information recorded in documents*', he is referring to a distinction that is fundamental to not just information retrieval, but to LIS at large. Vickery uses the example of a person asking about the height of Mount Everest. The answer can be either: '*The height of Mt. Everest is given on page 900, volume 8, of Encyclopaedia Britannica*' or it can be '*The height of Mt Everest is 9,002 ft.*' (Vickery 1961, pp. 2-3). While the first answer points the person to a source, the second one gives the questioning person a direct answer. Vickery continues that he '*restrict[s] the problem of Retrieval to the first type of information – the provision of the documents themselves, or the information about documents*' (p. 3). A source like the *Encyclopaedia Britannica* can be evaluated based on what we know *about* the source and its trustworthiness, but the content of a direct answer is harder to evaluate if you do not know the subject in advance. While it may sound bold to compare the *Encyclopaedia Britannica* with contemporary technologies, the comparison is conceptually intriguing, a point to which I will return throughout the paper.

The next quote is taken from an interview 62 years after the publication of Vickery's monograph and concerns the development of AI-infused search engines. When the CEO of Microsoft, Satya Nadella, was asked how the search engine Bing would be powered by ChatGPT, he replied, '*It's not just a search engine; it's an answer engine — because we've always had answers, but with these large models, the fidelity of the answers just gets so much better*' (Patel, 8 February 2023). Nadella refers to a change in Bing's focus from providing links to web objects to answering specific questions. What happens when AI-infused information systems increasingly provide answers rather than directing people to sources? When and how are such answers stabilised as sources and embedded into the information infrastructure – if at all? How might this shift in focus affect the disciplinary

understanding of certain core concepts in the field – specifically the concepts of search, source and evaluation?

In the next section, I ground my discussion in a long ongoing, disciplinary conversation on the materiality of information. I then proceed to explore the topic in the context of search, sources and evaluation. The paper concludes with a discussion on the potential for human agency within society's increasingly AI-infused, corporate information infrastructure.

Information in its material form

A starting point for this paper is materiality of information, its organisation within social institutions and information systems and the various practices through which people interact with information. As Paul Dourish (2007, p. 205f) suggests, a material understanding of information should be seen in contrast to a decontextualised virtual understanding. Other researchers have used related concepts to emphasise similar understandings of information. Marcia Bates (1999) refers to '*recorded information*', Mikael Buckland (1991) writes about '*information-as-thing*' (Buckland 1991), and more recently Vivian Petras (2023, p. 581) suggests that information science should be understood as '*concerned with how information is manifested across time and space*'. In my understanding of the materiality of information, material is not necessarily tangible. The concept of materiality encompasses the inherent contextuality of information: how information is produced, communicated and used in various practices; how the structures and processes of information systems (in the broadest sense) constrain and enable people's interactions with these systems; and how these interactions are shaped by historical, geographical and political institutions.

An interest in the materiality of information in LIS can be discussed in relation to the documentalist tradition, which has its roots in the writings of Paul Otlet from the late nineteenth and early twentieth centuries (e.g., Buckland 1997), and was further developed by Suzanne Briet's theorisation of documents in the mid-twentieth century (Day 2008). In her famous description of the boundaries of documents as a concept, Briet points out that an antelope can be understood as a document if it is placed in the zoo to be informative. In other words, an object is a document if it is '*processed for informational purpose*' (Buckland 1991, p. 355). According to Briet, a document is '*any concrete or symbolic indexical sign [indicie], preserved or recorded towards the end of representing, of reconstituting, or of proving a physical or intellectual phenomenon*' (Briet in Day 2008, p. 155). Documents in this sense can be seen as sources of information from which others can learn and gain knowledge, or in the words of Mikael Buckland (1991, p. 354), '*The documentalist's approach was to use 'document' as a generic term to denote any physical information resource rather than to limit it to text-bearing objects in specific physical media such as paper, papyrus, vellum, or microform*'.

W. Boyd Rayward (1991) has pointed out that Otlet was a pioneer in emphasising the social dimension of information, in particular by highlighting the role of information systems and institutions in promoting unity and world peace. However, Otlet's writings also contain ideas that can be discussed in relation to our contemporary information infrastructure. In the documentalist tradition, facts play a central role. According to Otlet, facts can be extracted from a document by a professional documentalist and form a new, highly standardised, short document. Otlet refers to this process as *documentation*. Bernd Frohmann (2008, p. 82) argues that Otlet, through the process of documentation, '*wants to eliminate authorial subjectivity*'. In the idealistic era in which he worked, Otlet wanted to create what he called a '*Universal book of knowledge*', with all possible facts extracted from books, journals, statistics and so on. These facts '*will be recorded on separate leaves or cards rather than being confined in volumes*' and organised according to a standardised classification system (Universal Decimal Classification) (Otlet in Frohmann 2008, p. 79). Otlet's idea of extracting facts from documents conveys an understanding of facts as abstract, neutral and

decontextualised from authors as well as from a specific volume (or similar). Can Otlet's positivist notion of extracting abstract facts from documents (or information sources) be compared to the increasing trend of information systems to attempt to provide answers instead of sources? I return to this comparison further on in the paper.

I would like to expand the understanding of materiality to include the theoretical notion of *sociomateriality*. Sociomateriality refers here to how information, information systems, people, imaginaries and practices are constantly configured and made together, always in motion and never stable, without human agency necessarily playing a more important role than non-human agency (e.g., Haider and Sundin, 2024). There are a number of scholars from various disciplines across the humanities and social sciences who can be related to a sociomaterial understanding (e.g., Dourish, 2017; Introna, 2013; Suchman, 2007). Dourish (2017, p. 3) writes '*that the social world manifests itself in the configuration and use of physical objects and that the properties of those physical objects and the materials from which they are made [...] condition the forms of social action that arise around them*'. A sociomaterial dimension of information emphasises the need to understand how agency is made in the constant configurations of information infrastructures and practices (Barad, 2007, p. 141).

How can a sociomaterial understanding of information help us comprehend the increasingly AI-infused information infrastructure from the perspective of LIS? I will now turn to the role of sources in information seeking through search engines.

From providing documents to providing answers?

As long as there have been such things as information systems, libraries and similar institutions for organising information, principles and methods for presenting and finding information have played a fundamental role. There are bibliographies, lists, card catalogues, book indexes, OPACs, classification systems, thesaurus with controlled vocabularies, web portals, bibliographical databases, reference management systems and so on. These systems organise and control access to information sources. New and old technologies often merge. Web search engines rely on their index, which only contains part of the web. This means that when we use Google Search, we are searching in an index, not the online web, and the size and freshness of the index are central quality criteria (Lewandowski, 2023, p. 41). These systems of information always bring with them a particular perspective, a view of the world, whether it is the representation of topics in library catalogues (Olson, 2002), web search engines (Lewandowski, 2023, p. 265; Noble, 2018) or generative AI (Sun et al., 2024).

Just as a librarian can provide answers to direct questions, there have always been exceptions to the main focus on search engines providing sources rather than answers. For example, Shah et al. (2009) discuss social Q&A systems and relate them to (and argue for their difference from) library reference services. Rich Gazan (2011, p. 2302) describes how the 1999 TREC conference included a track on the evaluation '*of systems designed to retrieve answers within documents, not just documents themselves, in response to test queries*'. Despite these examples of the opposite, search engines such as Google Search have traditionally provided users with a list of links to information sources (or documents). Google Search was introduced in 1998 as a search engine designed to provide users with links to documents. The '*ten blue links*' on the search engine results page (SERP) became a coveted goal for web producers seeking to appear among them. A major change happened in 2012 when Google Search started to incorporate results from the so-called Knowledge Graph into visual Knowledge Panels placed on the right-hand side of the SERP (Sullivan, 2020). If someone searched for the height of Mount Everest, as in the example of Vickery (1961) searching for factual information, the user could find the content directly from a Knowledge Panel, with text and images coming from open sources, without having to click on a link. However, the content in the Knowledge Graphs have identifiable sources. According to Google (2024): '*Facts in the*

Knowledge Graph come from a variety of sources that compile factual information. In addition to public sources, we license data to provide information such as sports scores, stock prices, and weather forecasts.'

Another significant development occurred in 2016, when Google Search began providing users with extracts from documents displayed above the organic search results in response to a query (Strzelecki and Rutecka, 2020). Through so-called Featured Snippets, users could, for example, access a pancake recipe directly on the SERP rather than just links to documents containing the recipe. Knowledge Panels, Featured Snippets and similar features have made sources less visible. At the same time, users such as the pupils quoted in the introduction to the paper are not primarily interested in finding sources – they are simply looking for a quick answer to their question. Generally, people tend to trust Google to such an extent that they attribute the answers they receive directly to Google itself (Pan et al., 2007; Schultheiß et al., 2018). Knowledge Panels, Featured Snippets, and other forms of direct answers provided by search engines exemplify a shift towards delivering content in the form of answers to questions rather than directing users to sources. When answers to user queries are presented directly on the SERP, the need to click on a link and identify a source becomes redundant.

The expansion of generative AI in recent years has driven the development from searching for documents to searching for answers even further. When Open AI launched ChatGPT in autumn 2022, public interest in generative AI was awakened. Suddenly, AI production of text, images, code and other outputs became possible for the general public in many countries, and social media has since then been flooded with synthetic information. As illustrated by the opening quote, LLM applications can be conceptualised as a replacement for search engines. The possibilities of generative AI appear limitless, and public discourse has been (and continues to be) characterised by significant hype. However, critical voices have also emerged. Among other concerns, criticism focuses on the risks that arise when users treat the output of generative AI applications in the same way as search engine results (e.g., Bender et al., 2021; Shah and Bender, 2022, 2024). There is a fundamental difference between using a search engine and an LLM application: while search engines traditionally provide links to websites and other documents, generative AI chatbots generate content without a stable source other than the LLM itself. Shah and Bender (2022, p. 8) note: '*Generative systems provide a new kind of possibility for interaction, namely conversational chat. However, this comes at the cost of direct access to sources.*' In contrast to Otlet's decontextualised and highly abstract fact leaves, the output of generative AI is devoid of retraceable connections to sources. Where Otlet's idea of a Universal book of knowledge depended on the labour of many human documentalists, texts generated by LLMs are instead the output of a '*stochastic parrot*' (Bender et al., 2021, p. 611) that, as expressed by Emily Bender (2024) relies on the technique of '*statistical models of the distribution of words in text, set up to output plausible-sounding sequences of words*'. This is a significant break with the prevailing LIS tradition of constructing systems that provide the user with a source of information.

A sociomaterial understanding of search and web search engines emphasises the need to include both an analysis of the practices in which information is searched, mediated and used, and an analysis of the technical side of how infrastructures constrain people's activities, and how they are made together. One practice in which searching for information with search engines has long played an important role is school education. Louise Limberg and her colleagues (e.g., 2008) showed in a series of studies a couple of decades ago that information search in Swedish schools often involves fact-oriented questions; what Gary Marconini (2006, p. 42) refers to as look up searches. Limberg et al. (2008; see also Rieh et al., 2016) have paid attention in their research to the role of search engines in co-producing the focus of fact-seeking in schools. If the evaluation and analysis of information is increasingly outsourced to AI, it is an important research task for the future to investigate how the formulation of school tasks could change together with AI.

Generative AI, such as applications based on LLM, produces information without stable links to sources, although – as touched on in the next section – there is also an interesting merge of search engines and AI technology through Retrieval Augmented Generation (RAG). From the user's point of view, the above brief history of search engines shows that the development of getting an answer instead of a link to a source when searching for information is not new. However, from a technical infrastructure point of view, a web search engine and a pure LLM application work very differently.

Below I develop my argument as to why this increasing invisibility of sources is important and likely consequential. I also approach the question of if and when answers to questions can be stabilised and become a source.

Evaluation of information

The launch of generative AI applications for the general public in 2022 posed a major challenge to the tradition of information literacy, particularly in relation to the evaluation of information sources. Programmes for information literacy training can be found in numerous examples online, such as the Association of College and Research Libraries (ACRL 2015) *Framework for information literacy for higher education*. The first frame out of six is titled '*Authority is constructed and contextual*'. The text under this heading begins with: '*Information resources reflect their creators' expertise and credibility, and are evaluated based on the information need and the context in which the information will be used*'. Although it is not explicitly stated, the ACRL framework seems to understand expertise as a human characteristic. The text continues: '*Authority is constructed in that various communities may recognise different types of authority. It is contextual in that the information need may help to determine the level of authority required.*' Wilson (1983, pp. 166-169), who has devoted an entire book to cognitive authority, argues that '*the obvious basis for recognising the cognitive authority of a text is the cognitive authority of its author*'. The author may be a person or an institution, but in Wilson's writing, a text always has an author with an intention. Wilson also suggests two other factors that should be considered when assessing the cognitive authority: '*publication history*', which refers to the source of a text, and '*intrinsic plausibility*', which refers to how readers evaluate the content in relation to their previous experiences and knowledge. A significant challenge in the evaluation of information lies not only in the increasing invisibility of sources on search engine results pages (SERPs) or even the disappearance of traceable sources in generative AI, but also in the shifting understanding of what constitutes an originator – or, in the context of the ACRL, a creator. This shift fundamentally alters how we can assess the credibility, accountability and authenticity of information in the digital landscape. In relation to Wilson's (1983) argument, AI-generated content cannot be evaluated by assessing the author or publication history. Instead, we are left with *intrinsic plausibility*.

Most information literacy guides, models and methods include a discussion of the importance of analysing the originator of information: who the originator is, what the originator has previously published and so on. Another example is the SIFT (Stop, Investigate, Find, Trace) method, developed by digital literacy researcher Mike Caulfield. The method aims to train the user to ask the following questions in relation to the originator:

> *If the creator is a media outlet, are they reputable and respected, with a recognised commitment to verified, independent journalism?*
>
> *If it's an individual, what expertise do they have in the subject at hand (if any)? What financial ties, political leanings or personal biases may be at play?*
>
> *If it's an organisation or a business, what is their purpose? What do they advocate for, or sell? Where does their funding come from? What political leanings have they shown?* (Ruggeri, 2024)

Who is the originator of a ChatGPT output? As Bender et al. (2021) emphasise, is the generated text of an LLM produced by a '*stochastic parrot*', without links to a stable source and without an identifiable originator. In his well-known essay 'What is an author?' Foucault (1969) challenges the notion of the unique individual author. Instead, he advances the concept author function. The analytical comparison of Foucault's concept with the functions of generative AI is perhaps daring. Yet, the generative AI bots that we encounter do place a function rather than an author or other originators at the centre of their varying outputs. In the case of synthetic information, the function of the originator manifests as a configuration of an AI model, encompassing its algorithms, training data, fine-tuning processes, and – critically – the user prompts carried out within sociomaterial practices. What emerges is not the intention of an originator in the traditional sense, but rather a composite of the intentions of the LLM developers and the prompter – appropriating; perhaps exploiting – content produced by others, for others, at different times and for different purposes. Again, sources are becoming increasingly invisible or – in the case of LLM chatbots – nonexistent, making the evaluation of information quality exceptionally challenging.

A hybrid form of generative AI and web search is based on RAG technology. RAG integrates AI functionality into search applications such as Perplexity, Google Overview, and Bing Copilot, enabling a chatbot to summarise the content of top-ranked links. Although links to sources are provided, the summaries are presented in such an authoritative manner that it is reasonable to assume many users will find the summary sufficient and forego consulting the original sources. Google even states, '*Let Google do the searching for you*' (Reid, 14 May 2024). There are examples of how such summaries can lead to the spread of conspiracy theories, for example about the climate crisis (Haider, 27 August 2024) or '*scientific racism*' (Gilbert, 24 October 2024). Traditionally, a web search engine provides the user with sources via links that the algorithm interprets as relevant, and people tend to trust search engine rankings. AI-infused search engines based on RAG technology outsource not only the selection and ranking of sources but also their evaluation. While users are accustomed to relying on Google (or other search engines) to perform the searching, they are not accustomed to these platforms taking on the '*sense-making*' (Bender 2024) of content on their behalf. Haider (27 August 2024) describes this outsourcing of evaluation as an '*automation of media and information literacy*'. The quality of the summaries generated by these systems depends heavily on the quality of the sources identified by the search engine. In the examples by Haider (27 August 2024) and Gilbert (24 October 2024), the summaries became problematic because they were derived from low-quality sources. This issue parallels the problem of data voids (Golebiewski and boyd, 2018); however, the increasing invisibility of sources to users exacerbates the severity of the problem. For topics with few sources or sources of questionable quality, these voids can be exploited to amplify the spread of conspiracy theories and other forms of disinformation.

If users outsource their evaluation of sources to an AI-powered web search engine, as in the case of RAG, how do search engines evaluate synthetic information? According to Google, the search engine treats human-generated information no differently to synthetic information. The company argues that it focuses on the quality rather than the producer: '*Our focus on the quality of content, rather than how content is produced, is a useful guide that has helped us deliver reliable, high-quality results to users for years*' (Google, 8 February 2023). In a way, the idea of content without an identifiable originator can be related to Otlet's idea of extracting pure facts from documents and categorising them into a new context according to a classification system. When AI-generated output is incorporated into a technical manual, a social media post, a newspaper article or a school essay, the text is stabilised by a more traditional originator – such as a company, a newspaper and its journalists, a student or another individual – who then becomes accountable for the content. Just as decontextualisation and recontextualisation take place in documentation, if documents include AI-generated content, they are rearranged by the ranking of the sources in the search

engines. The content is assigned a source and search engine contributes to stabilising the source of the AI-generated content.

Conclusion

As Wilson (1983) reminds us, our understanding of the world depends on what '*others*' tell us. What if the '*others*' are AI? The pupils in the opening quote like ChatGPT because the application gives a '*more direct answer*' (Wasell, 23 June 2023). At the same time, LIS as a discipline – with concepts such as search, source and evaluation – emerged primarily from a tradition of documents, and systems for finding documents (including digital documents, of course). Vickery's distinction between finding '*information about documents*' and finding '*information recorded in documents*' plays an important role as a reference throughout the paper. Developments of Google Search (e.g., Knowledge Panels and Featured Snippets) have increasingly led to the answers to users' questions – the content – appearing directly on the SERP, as a complement to the prevailing LIS tradition, which has (with exceptions, of course) traditionally been about understanding how systems provide users with sources of information. As people begin to also use chatbots based on LLMs to search for (or rather produce) information, content is replacing sources. From a user perspective, this evolution appears gradual, if noticed at all; however, from a technical standpoint, the shift is dramatic. The link between content and sources is broken. Simultaneously, web search engines and AI are increasingly being integrated through RAG technology, which enables first the creation of search terms and then the summarisation of content from the top links on the SERP. I argue that the growing prevalence of generative AI applications based on large language models (LLMs), along with the integration of LLMs into search systems, is shifting the focus towards searching for content rather than identifying sources.

Increasing inbuilt AI support for searching and drawing conclusions from information sources may lead to de-skilling of searching for and evaluating information. Such de-skilling of information literacy transfers some of the previously required analytical skills from the user to the system. As a result, this transfer can also be seen as a transfer of agency in which the user has less control and becomes increasingly more dependent on the functions of the system – not only in the search as such, but also in evaluating and analysing the results. When evaluating and analysing generative AI output, verifying the originator becomes meaningless, as there is no originator with an intention, at least not in the traditional sense. But does it matter? Perhaps the focus should instead be on understanding how a LLM is trained, the role of fine-tuning in its moderation and how users formulate prompts? Prompting represents a co-construction of content, where the '*author function*' emerges through the interplay of users and AI systems within sociomaterial practices. However, once AI-generated content is incorporated and published somewhere, it becomes stabilised by its attribution to an originator's intent. As described earlier in the paper, stabilisation can be further reinforced by its inclusion in a search engine index.

This paper contributes to the conceptual discussion of searching, sources and evaluation within an increasingly AI-infused information infrastructure. Engaging in this discussion has prompted me to consider how meaningful conceptual contributions can be made when dealing with an information infrastructure that evolves so rapidly. The timeframe associated with academic writing, peer review, conferences and publishing is undoubtedly slower than the pace of technological change. My approach to addressing this challenge involves anchoring the discussion in the disciplinary tradition of LIS, and framing it through the overarching theoretical lens of sociomateriality. LIS, as a discipline, has decades of experience discussing information and researching information systems and their societal use. An important task for the discipline for the future is to create and communicate an understanding of the changing digital infrastructure, its impact on the role of sources and evaluation, and the increasing shift towards the search for answers rather than sources, all grounded in the analysis of specific practices.

Acknowledgements

About the author

Olof Sundin is Professor in Information Studies at Lund University. His research is concerned with the contemporary configuration and control of information and its societal implications. He can be contacted at olof.sundin@kultur.lu.se

References

ACRL (2015). *Framework for information literacy for higher education.* https://web.archive.org/web/20241122002148/https://www.ala.org/sites/default/files/acrl/content/issues/infolit/Framework_ILHE.pdf

Barad, K. (2007). *Meeting the universe halfway: quantum physics and the entanglement of matter and meaning.* Duke University Press.

Bates, M. J. (1999). The invisible substrate of information science. *Journal of the American Society for Information Science,* 50(12), 1043–1050. https://doi.org/10.1002/(SICI)1097-4571(1999)50:12<1043::AID-ASI1>3.0.CO;2-X

Bender, E. (2024). Information literacy and chatbots as search. In *Mystery AI hype theater 3000.* https://web.archive.org/web/20241120040030/https://buttondown.com/maiht3k/archive/information-literacy-and-chatbots-as-search/

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: can language models be too big? 🦜. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610–623). https://doi.org/10.1145/3442188.3445922

Buckland, M. K. (1991). Information as thing. *Journal of the American Society for Information Science,* 42(5), 351–360. https://doi.org/10.1002/(SICI)1097-4571(199106)42:5<351::AID-ASI5>3.0.CO;2-3

Buckland, M. (1997). What is a 'document'? *Journal of the American Society for Information Science,* 48: 804–809. https://doi.org/10.1002/(SICI)1097-4571(199709)48:9<804::AID-ASI5>3.0.CO;2-V

Day, R. E. (2008). 'A necessity of our time': documents and culture in Suzanne Briet's Qu 'est-ce que la documentation. In W. B. Rayward (Ed.), *European modernism and the information society: informing the present, understanding the past* (155–164). Ashgate.

Dourish, P. (2022). *The stuff of bits: An essay on the materialities of information.* MIT Press.

Foucault, M. (1971). What is an author? In D. F. Bouchard (Ed.), *Language, counter-memory, practice: selected essays and interviews.* Basil Blackwell.

Frohmann, B. (2008). The role of facts in Paul Otlet's modernist project of documentation. In W. B. Rayward (Ed.), *European modernism and the information society: informing the present, understanding the past* (pp. 75-88). Ashgate.

Gazan, R. (2011). Social Q&A. *Journal of the American Society for Information Science and Technology*, 62(12), 2301–2312. https://doi.org/10.1002/asi.21562

Gilbert, D. (24 October 2024). Google, Microsoft, and Perplexity are promoting scientific racism in search results. *Wired.* https://web.archive.org/web/20241123030003/https://www.wired.com/story/google-microsoft-perplexity-scientific-racism-search-results-ai/

Golebiewski, M. & Boyd, D. (2018). Data voids: where missing data can easily be exploited. *Data & Society.* https://web.archive.org/web/20240826040959/https://datasociety.net/library/data-voids-where-missing-data-can-easily-be-exploited/

Google (8 February 2023). *Google Search's guidance about AI-generated content.* https://web.archive.org/web/20240217024506/https://developers.google.com/search/blog/2023/02/google-search-and-ai-content?ref=404media.co

Haider, J. (27 August 2024). AI-powered web search gives climate deniers and conspiracy theorists free rein. *Tech Policy Press.* https://web.archive.org/web/20241107135044/https://www.techpolicy.press/ai-powered-web-search-gives-climate-deniers-and-conspiracy-theorists-free-rein/

Haider, H. & Sundin, O. (2024). Sociomateriality. In A. Hicks, A. Lloyd & O. Pilerot (Eds.), *Information literacy through theory* (pp. 149–164). Facet Publishing.

Introna, L. D. (2013). Epilogue: Performativity and the becoming of sociomaterial assemblages. In *Materiality and space: organizations, artefacts and practices* (pp. 330–342). Palgrave Macmillan UK.

Lewandowski, D. (2023). *Understanding search engines.* Springer.

Limberg, L., Alexandersson, M., Lantz-Andersson, A., & Folkesson, L. (2008). What matters? Shaping meaningful learning through teaching information literacy. *Libri*, 8, 82–91. https://doi.org/10.1515/libr.2008.010

Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4), 41–46. https://doi.org/10.1145/1121949.1121979

Noble, S. U. (2018). *Algorithms of oppression: how search engines reinforce racism.* New York University Press.

Olson, H. (2002). *The power to name: locating the limits of subject representation in libraries.* Springer.

Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: users' decisions on rank, position, and relevance. *Journal of computer-mediated communication*, 12(3), 801–823. https://doi.org/10.1111/j.1083-6101.2007.00351.x

Patel, N. (8 February 2023). Microsoft thinks AI can beat Google at search — CEO Satya Nadella explains why. *The Verge.* https://web.archive.org/web/20241119223442/https://www.theverge.com/23589994/microsoft-ceo-satya-nadella-bing-chatgpt-google-search-ai

Petras, V. (2023). The identity of information science. *Journal of Documentation*, 80(3), 579–596. https://doi.org/10.1108/JD-04-2023-0074

Rayward, W. B. (1991). The case of Paul Otlet, pioneer of information science, internationalist, visionary: reflections on biography. *Journal of Librarianship and Information Science*, 23(3), 135–145. https://doi.org/10.1177/09610006910230030

Reid, L. (14 May 2024). Generative AI in Search: let Google do the searching for you. *The Keyword*. https://web.archive.org/web/20241122105549/https://blog.google/products/search/generative-ai-google-search-may-2024/

Rieh, S. Y., Collins-Thompson, K., Hansen, P., & Lee, H. J. (2016). Towards searching as a learning process: a review of current perspectives and future directions. *Journal of Information Science*, 42(1), 19-34. https://doi.org/10.1177/0165551515615841

Ruggeri, A. (2024). *The 'Sift' strategy: a four-step method for spotting misinformation*. https://web.archive.org/web/20241126162247/https://www.bbc.com/future/article/20240509-the-sift-strategy-a-four-step-method-for-spotting-misinformation

Schultheiß, S., Sünkler, S., & Lewandowski, D. (2018). We still trust in Google, but less than 10 years ago: an eye-tracking study *Information Research*, 23(3), paper 799. http://InformationR.net/ir/23-3/paper799.html (http://www.webcitation.org/72M895A36)

Shah, C., & Bender, E. M. (2022). Situating search. In *Proceedings of the 2022 Conference on Human Information Interaction and Retrieval* (pp. 221–232). https://doi.org/10.1145/3498366.3505816

Shah, C., & Bender, E. M. (2024). Envisioning information access systems: what makes for good tools and a healthy Web? *ACM Transactions on the Web*, 18(3), 1–24. https://doi.org/10.1145/3649468

Shah, C., Oh, S., & Oh, J. S. (2009). Research agenda for social Q&A. *Library & Information Science Research*, 31(4), 205–209. https://doi.org/10.1016/j.lisr.2009.07.006

Strzelecki, A., & Rutecka, P. (2020). Direct answers in Google search results. *IEEE Access*, 8, 103642–103654. http://doi.org/10.1109/ACCESS.2020.2999160

Suchman, L. (2007). *Human-machine reconfigurations: plans and situated actions*. Cambridge University Press.

Sullivan, D. (2020). A reintroduction to our Knowledge Graph and knowledge panels. *The Keyword*. https://blog.google/products/search/about-knowledge-graph-and-knowledge-panels/

Sun, L., Wei, M., Sun, Y., Suh, Y. J., Shen, L., & Yang, S. (2024). Smiling women pitching down: auditing representational and presentational gender biases in image-generative AI. *Journal of Computer-Mediated Communication*, 29(1), zmad045. https://doi.org/10.1093/jcmc/zmad045

Vickery, B. C. (1961). *On retrieval system theory*. Butterworths.

Wasell, C. (June 13 2023). Nya spökskrivaren ChatGPT väcker blandade känslor i skolan – 'Den förändrar allt'. *Dagens Nyheter*. https://web.archive.org/web/20240213235619/https://www.dn.se/sverige/nya-spokskrivaren-chatgpt-vacker-blandade-kanslor-i-skolan-den-forandrar-allt/

Wilson, P. (1983). *Second-hand knowledge: an inquiry into cognitive authority*. Greenwood Press.