

Ali Shiri
University of Alberta, Edmonton, Alberta, Canada

CONCEPTUALIZATIONS OF INFORMATION SCIENCE BY LARGE LANGUAGE MODELS

Abstract: This paper reports a comparative study of how large language models understand and represent the domain of information science. Five large language models were selected for this study, namely ChatGPT, Perplexity.ai, Google Gemini, Meta AI and Claude. A set of five prompts was utilized in this study for comparison. The findings suggest differences and variations in how these LLMs conceptualize and represent information science, its definitions, and interdisciplinarity, theoretical models, and methods.

Introduction

Generative AI in general and large language models in particular present new and unprecedented opportunities and challenges for academic research and scholarship. Information science as a data and information-focused discipline, has a long history of addressing technical, social, physical, cognitive, behavioural, cultural, and ethical implications of data and information work. With the growing number of foundation models and large language models, academic research is beginning to experience a paradigm shift. Simply put, large language models and multimodal generative AI applications can be utilized to generate and brainstorm ideas and research topics; formulate hypotheses; design experiments; conduct literature review and synthesis (Wang et al., 2024); code, analyze, and visualize qualitative and quantitative data; develop methodological and theoretical frameworks; recommend sources and citations; summarize text and write abstracts; and assist in research dissemination and presentation. Considering the emergence of these new large language models and their potential to support research, the objective of this paper is to report on a study that explores the ways in which popular large language models conceptualize and represent information science as a domain. The key question in this study is how well large language models understand and represent the domain of information science. As large language models rely on larger datasets and new training models, one of the key questions is how they can be used for domain and discipline analysis and more broadly for knowledge representation. In other words, do the publicly available LLMs represent information science with similar traditional conceptual, domain, and disciplinary boundaries as traditional research? Of course, this study aims to scratch the surface and does not intend to compare LLMs performances with citation-based and bibliometric domain analyses. But it provides a critical perspective of how LLMs could be used for research and analysis purposes.

While LLMs have demonstrated powerful capabilities in providing plausible and human-sounding outputs to various types of prompts, there are significant concerns and challenges surrounding the critical, appropriate, and ethical use of generative AI tools in the responsible conduct of information science research. These include misrepresentation of data and authorship, difficulty in replication of research results, data and algorithmic biases and inaccuracies, user and data privacy and confidentiality, quality of outputs, data and citation fabrication, and copyright and intellectual property infringement. An analysis of 950,965 papers published between Jan 2020 - Feb 2024 on the arXiv, bioRxiv, and Nature portfolio journals found that 17.5% of computer science papers and 16.9% of peer review text had at least some content drafted by AI

(Liang et al., 2024a, 2024b). A recent study exploring the quality of a select number of LLMs (GPT, Claude, Google Gemini) in citing relevant medical references found that the models performed significantly worse on questions sourced from Reddit r/AskDocs versus either MayoClinic or UpToDate, two reliable and widely known sources of medical advice for physicians (Wu et al., 2024). It also revealed that the cited URLs are from US-based sources over 90% of the time, potentially reflecting American patient-centric medical evidence and standard of care. In another study of the repeatability, reproducibility, and accuracy of ChatGPT to perform emergency department triage, the researchers found that ChatGPT was not sufficient for emergency physicians to triage simulated patients. They concluded that ChatGPT showed poor repeatability when the same prompt was used repeatedly, and the overall accuracy of triage was only 47% (Franc et al., 2024). In this study, we are interested in how a select number of LLMs conceptualize and represent the domain of information science and its definitions, theories, and methods. For clarity the terms domain and discipline have been used here interchangeably. It should be noted that there are implied underlying factors such as the deficiencies in training data and the ways in which these platforms generate responses. This is an exploratory study to examine how an example domain, such as information science is represented in a select number of large language models.

Context

Following the introduction of ChatGPT in November 2022, Large Language Models (LLM) have gained increasing popularity and prominence. LLMs are considered a category of Pre-trained language models that are large-scale in terms of both model and data. They increasingly exhibit surprising and powerful capabilities known as *emergent abilities*, namely in-context training, instruction following, and step-by-step reasoning (Zhao et al., 2023). The notion of scale in pre-trained language models is particularly important as it implies significantly larger sizes, extensive training data, and self-supervision in training and learning from data (Liu et al., 2024). Popular and widely used examples of LLMs include ChatGPT, Google Gemini, Meta AI, and Claude. Considering the diversity and scale of the data used for pre-training, these LLMs provide varying levels of performance and functionality. For instance, ChatGPT makes use of a wide range of training data sources such as books, social media platforms, Wikipedia, news articles, websites, code repositories, forums, academic research papers, and public web content.

One of the heated debates concerning large language models is whether they are capable of understanding data, concepts, and contexts (Mitchell, M., & Krakauer, 2023). One key argument asserts that large language models cannot understand data and concepts as these models follow statistical correlations and machine learning patterns. Other researchers argue that the LLMs have the potential to teach us as humans about language understanding and personhood (y Arcas, 2022). Recent research on the effectiveness of LLMs for domain-specific and interdisciplinary research tasks has found that LLMs can have great potential in such areas as biological sciences, chemical sciences, engineering, environmental as well as social sciences and can contribute to the elimination of disciplinary silos and enhanced interdisciplinarity in research (Boyko et al., 2023). Considering the emerging nature of large language models, there is very little research in understanding how LLMs understand and conceptualize academic disciplines and their boundaries.

This paper aims to use information science as a case to conduct a preliminary exploration of how LLMs understand and conceptualize information science as a domain. The conceptualizations and definitions of information science have been extensively discussed in several studies. Here we provide some examples to show the extent of the literature on definitions and domain mapping of information science. Over the past seven decades, numerous researchers have explored and provided definitions and conceptualizations for information science and its nature, scope, and identity (Borko, 1968; Belkin, 1978; Bates, 1999, Petras, 2024). Numerous studies have been conducted to explore domain analysis, domain mapping, and bibliometric knowledge mapping of information science, aiming to capture the boundaries and scope of information science as a domain (Hjørland, 2002; Zins, 2007; Janssens et al., 2008; Chua, A. Y., & Yang, 2008; Larivière et al., 2012; Xu et al., 2016; Wijewickrema, 2023).

While the above studies help us understand the history, definitions, and thematic characterization of information science, this study specifically aims to explore the outputs of a select number of prompts related to the conceptualizations of information science as a discipline in LLMs. Particular emphasis is placed on the definitions of information science as a discipline, and its theoretical and methodological characteristics.

Methods

To conduct this study, five research questions were developed as prompts to be submitted to a select number of LLMs. In formulating these questions, we considered the historical context of information science, its interdisciplinary nature, and the arguments that have appeared in the literature on whether or not information science is a distinct discipline with its own theories and methodologies. The following are the questions/prompts that were used in this study:

Prompts

The following prompts were created by the author without any assistance from large language models. This was an intentional decision in order to create a degree of consistency and predictability.

1. What is Information Science? Suggest a definition
2. Is information science a domain?
3. Is information science a discipline or an interdisciplinary domain? if it is interdisciplinary, what are the main disciplines in information science?
4. Does information science have a theory or theories? Suggest important and influential theoretical frameworks in Information Science
5. What research methods are used in information science?

You will notice that there are similar and synonymous concepts and connotations in the above prompts. The decision to use various conceptually similar terms such as domain and discipline and theories and theoretical is deliberate as LLMs have been found to produce inconsistent and different outputs related to the same or conceptually similar prompts. Table 1 shows the five large language models selected for this study. The selection criteria simply included availability, reliable access, and relative popularity. Comparative performance analysis of these models has

been conducted in many different studies and disciplines including medicine, business, law, and the sciences. These models exhibit different functionality in terms of speed, quality, reliability, and repeatability of outputs. However, they are among the most widely used LLMs around the world. Some of these LLMs identify themselves as search engines or Retrieval Augmented Generation (RAG) platforms, including Perplexity.ai and Google Gemini. Retrieval Augmented Generation platforms retrieve relevant external information (e.g., documents, databases, or APIs) to enhance their ability to generate accurate and context-aware responses. One of the issues with large language models that do not support RAG is that their outputs may not include the most recent and emerging trends and developments.

The above prompts were simultaneously input in the five LLMs on December 5 and 6, 2024.

Table 1. Selected large language models

LLM model/version	Underlying LLM	Release date
ChatGPT 40 (subscription version)	ChatGPT	May 2024
Perplexity.ai (Free version)	ChatGPT (40) & Claude (3.5)	December 2022
Google Gemini 1.5 (Free version)	Gemini model	May 2024
Meta AI (Free version)	LLaMA 3.1	April 2024
Claude 3.5 Haiku (mobile, free version)	Claude 3 Haiku	October 2024

Preliminary Analysis

A table was created to analyze and compare the outputs from the five LLMs in response to the five research questions. A preliminary analysis of the data is presented below based on the research questions.

Question 1. What is Information Science? Suggest a definition

The objective of the first simple question was to examine how the LLMs view information science. ChatGPT, Claude, and Google Gemini consider information science to be interdisciplinary. Meta AI uses the term multidisciplinary and Perplexity.ai does not use any of these terms in its definition. All of the LLMs use similar concepts and terms to define information science, including collection, organization, representation, storage, retrieval, and dissemination of information. Google Gemini seems to rely on Borko's (1968) definition of information science emphasizing 'the study of the properties and behavior of information,

including its creation, collection, organization, storage, retrieval, interpretation, transmission, transformation, and utilization’. An interesting observation here is that Perplexity.ai is the only platform that provides a list of citations for transparency, including such sources as Wikipedia, Britannica, and the University of Michigan School of Information sources. Other LLMs do not provide such references unless they are specifically prompted to do so. Some of the LLMs provide a more detailed description of the specific aspects of information science as shown in Table 2. One noteworthy observation is that while there are commonalities among these definitions, some tend to be more comprehensive than others, highlighting the fact that they use a variety of sources to train models. This is particularly important as explainability and transparency are key principles how LLMs are developed and deployed.

Table 2. Comparison of information science definitions in LLMs

Prompt	ChatGPT	Perplexity.ai	Google Gemini	Meta AI	Claude
What is Information Science? suggest a definition	This field focuses on optimizing information systems and processes to enhance access to knowledge, support decision-making, and drive innovation across diverse domains, including academia, business, healthcare, and technology.	It encompasses the analysis of information processes and the development of technologies that facilitate effective information handling across various domains.	Information Science is the interdisciplinary study of the properties and behavior of information, including its creation, collection, organization, storage, retrieval, interpretation, transmission, transformation, and utilization. It explores the relationships between people, organizations, and information systems, with the goal of improving access to and understanding of information.	Information Science is the systematic study and application of concepts, theories, and technologies related to the creation, collection, organization, storage, retrieval, dissemination, and utilization of information, with the aim of enhancing knowledge, decision-making, and innovation in various disciplines and domains.	It integrates knowledge from various disciplines such as computer science, library science, cognitive science, sociology, and communication studies to understand how information is generated, processed, and utilized in human and technological contexts.

Question 2. Is information science a domain?

Considering the evolving, constantly changing, and interdisciplinary nature of information science, this question aimed to invite arguments for and against the nature of information science as a domain. All the LLMs responded by noting that information science is indeed a domain and is an interdisciplinary field. Perplexity.ai and Claude specifically stressed that information

science is an academic field of study. As can be seen from Table 3, the LLMs in this study emphasize different aspects of the concept of information science. For instance, ChatGPT provides a more nuanced understanding of information science to include the elements of technology, human behaviour, and organizational contexts, whereas Google Gemini lists the intersection of information science with other disciplines and domains. Claude mainly offers a standard definition of information science.

Table 3. Is information science a domain?

Question	ChatGPT	Perplexity.ai	Google Gemini	Meta AI	Claude
Is information science a domain?	Yes, information science is a domain. It is an interdisciplinary field that focuses on the study, design, management, and use of information systems and processes. It encompasses a broad scope, integrating elements of technology, human behavior, and organizational contexts to address how information is created, organized, accessed, and used.	Yes, information science is indeed a domain. It is an interdisciplinary academic field and profession with a distinct scope and focus	Yes, information science is indeed a domain. It's an interdisciplinary field that studies the collection, organization, dissemination, interpretation, and use of information. This domain intersects with various disciplines, including computer science, library science, cognitive science, and social sciences.	Yes, information science is indeed a domain. It is an interdisciplinary field that deals with the study of information systems, including the storage, retrieval, and dissemination of information.	Yes, information science is indeed a recognized academic and professional domain. It is an interdisciplinary field that focuses on the collection, classification, manipulation, storage, retrieval, movement, dissemination, and protection of information.

Question 3. Is information science a discipline or an interdisciplinary domain? if it is interdisciplinary, what are the main disciplines in information science?

The rationale for this question was to explore whether there are any differences and similarities among the LLMs in listing the disciplines within information science. Once again, there is consistency in LLMs in noting that information science is an interdisciplinary field. ChatGPT and Perplexity.ai mention that information science is an interdisciplinary domain and not a single discipline. Claude notes that information science is fundamentally an interdisciplinary domain, “though it has evolved to develop its own distinct identity and core body of knowledge”. All the LLMs note that information science integrates, combines, and draws upon concepts, theories, methodologies, applications, and approaches from various disciplines. Table 4 shows similar

disciplines within information science as mentioned by all of the LLMs, namely Library Science, Cognitive Science, Computer Science, and Sociology. Four out of the five LLMs list ‘Linguistics’ as a related discipline. ChatGPT provides an additional delineation by noting that due to the interdisciplinary nature of information science, it serves as a bridge between these disciplines, combining their methodologies and insights to address complex problems related to information in diverse contexts and includes Digital Libraries, Data Science, Health Informatics, and Artificial intelligence as examples. This is particularly interesting since these research and development areas represent closer ties between Information Science and Computer Science. Perplexity.ai and ChatGPT are the only LLMs that list ‘Law’ as a related discipline. Of interesting note is also that Claude and Perplexity.ai are the only LLMs to mention ‘archival science’ and ‘archival studies’ as a related discipline.

Table 4. Disciplines in information science

Question	ChatGPT	Perplexity.ai	Google Gemini	Meta AI	Claude
Is information science a discipline or an interdisciplinary domain? if it is interdisciplinary, what are the main disciplines in information science?	<ul style="list-style-type: none"> -Library Science -Cognitive Science -Philosophy of information Technical disciplines -computer science -mathematics and statistics -Linguistics Applied and social sciences -Communication studies -Management science -Sociology -Education -Law and Policy studies -Geography Digital Libraries Data Science Health Informatics Artificial intelligence 	<ul style="list-style-type: none"> -Computer science -library science -Cognitive science -communication -management -Mathematics and statistics -Social sciences -Law -Philosophy -Archival Science 	<ul style="list-style-type: none"> -Computer science -Library science -cognitive science -social sciences -mathematics -Linguistics 	<ul style="list-style-type: none"> -Computer science -Library and Information Studies (LIS): -Cognitive Science -Linguistics -Sociology -Psychology -Education -Communication Studies Management and Organization Studies 	<ul style="list-style-type: none"> -Computer science -cognitive science -Library and archival sciences -sociology -communication studies -linguistics -anthropology -Management and Organizational Studies

Question 4. Does information science have a theory or theories? Suggest important and influential theoretical frameworks in Information Science

This question as well as question 5 aim to address a nuanced understanding of how information science is understood by LLMs, particularly from the perspective of theoretical and methodological frameworks and contributions. All five LLMs note that information science has

theories, but emphasize that there is no overarching (Meta AI), universally agreed upon (ChatGPT) theory as it is a young (Google Gemini) and interdisciplinary (Perplexity.ai) domain that lacks a single comprehensive unified theory (Claude). While ChatGPT provides a more comprehensive list, none of the LLMs provides a comprehensive list of the theories developed and used in Information Science. However, there are several theories that are common in all or most of the LLMs. These include Cognitive theory, Information Theory, Social Theories, and Information Behaviour Theory. Dervin's sense-making theory is mentioned by ChatGPT and Claude. ChatGPT also lists theories by Chatman and Wilson. A quick glance at Table 5 shows that Meta AI provides an interesting information retrieval-focused list of theories that are not included in other models. Together these different LLMs provide a more comprehensive overview of the theoretical frameworks and models both developed and used in information science.

Table 5. Information Science theory/theories

Question	ChatGPT	Perplexity.ai	Google Gemini	Meta AI	Claude
Does information science have a theory or theories?	<ul style="list-style-type: none"> -Information Behavior Theories (Chatman, Wilson) -Cognitive and Psychological Theories (sense-making by Dervin) -Systems Theory -Information Theory (Mathematical) -Critical and Social Theories -Knowledge Organization Theories -Information Ecology -Data-Information-Knowledge-Wisdom (DIKW) Model -Network Theory -Ethics of Information 	<p>Information science incorporates theories from: - information theory, -coding theory, -cognitive theory, -social theories, -organizational theories, - systems theories,</p> <p>Emerging Theoretical Frameworks</p> <p>-Human information Coordinating behaviour that comes from an article by Amanda Spink in 2000</p> <p>-Holistic models</p>	<ul style="list-style-type: none"> -Information Theory -Cognitive Information Processing -Use and Gratification - Theory Social Construction of Technology (SCOT) -Diffusion of Innovation Theory 	<ul style="list-style-type: none"> -Information Retrieval Theories (boolean, Vector space, probabilistic) -Information Behavior Theories (sense-making, - Information search process) -Information Systems Theories -Knowledge Management Theories -Human-Computer Interaction (HCI) Theories 	<ul style="list-style-type: none"> -Shannon-Weaver Information Theory -Cognitive Information Theory -Social Information Theory -Sense-Making Theory -Ecological Information Theory

The responses to the second part of this question, namely 'Suggest important and influential theoretical frameworks in Information Science', were particularly interesting. ChatGPT provided a more detailed list, specifically referencing Carol Kuhlthau (Information Search Process),

Brenda Dervin (sense-making), Tom Wilson (Information Behaviour Model) Elfreda Chatman (Theories of information behaviour), Karen Fisher (Information Grounds), and Reijo Savolainen (Everyday information seeking model). Perplexity.ai referred to the Framework for Information Literacy in Higher Education by the American Library Association. Claude also provided a more detailed list of theories, including Information processing theory, Sense-making theory, Actor-network theory, Information Ecology Framework, Social Construction of Technology, Kuhlthau's model of the Information Search process, and Systems theory. Google Gemini lists Information theory, Cognitive information processing, Network theory, and Big data analytics. Meta AI tends to focus on information retrieval models, information behaviour theories (sense-making is the only one listed), knowledge management theories, and human-computer interaction.

Of course, the theories that these LLMs suggest are not the only ones in information science. For instance, relevance is one of the foundational concepts and constructs in information science that has been widely researched as a theory that is shared with the information retrieval community and represents one of the key components of several information search and interaction behaviour models. The LLMs used in this study fail to provide deep and nuanced understanding of such concepts that contribute to our understanding of the discipline.

Question 5. What research methods are used in information science?

A comparison of commonalities and differences in the outputs from the five LLMs for this prompt is provided in this section (Table 6). In general, Qualitative methods, Quantitative methods, Mixed methods, and Emerging and Specialized methods were suggested by all of the LLMs. For instance, all models mention core qualitative methods such as interviews, focus groups, case studies, ethnography, content analysis, and discourse analysis. Quantitative methods such as surveys, experiments, statistical analysis, and bibliometric methods are universally noted. The models also stress the importance of combining qualitative and quantitative approaches for comprehensive insights. Most of the LLMs discuss methods such as digital tools (text mining, NLP, and social media analysis) and practical approaches such as usability studies or participatory design. There are differences in how these LLMs view research methods in information science. For instance, ChatGPT provides a broader range of methods, including participatory/action research and computational/digital methods (e.g., network visualization, web analytics). Perplexity.ai emphasizes data collection techniques such as direct/indirect observation and archival record analysis. While Google Gemini suggests specific mixed-methods strategies such as combining qualitative and quantitative techniques for problem-solving, it does not provide an elaborate account of how these could be achieved. Meta AI specifically mentions action research, design science research, and discourse analysis and introduces "practice-based research" focusing on bridging theoretical and applied aspects. Claude suggests computational methods (network analysis, machine learning) and highlights critical research methods that address power structures and social justice. While ChatGPT provides the most expansive categorization of methods used in information science integrating traditional and modern methods, Perplexity.ai presents a richer treatment of observational and data collection techniques.

Table 6. Information science research methods

Prompt	ChatGPT	Perplexity.ai	Google Gemini	Meta AI	Claude
What research methods are used in information science?	<p>Qualitative methods: interviews, focus groups, ethnography, case studies, content analysis, grounded theory, discourse analysis</p> <p>-Quantitative methods: Surveys and Questionnaires</p> <p>-Bibliometric and Scientometric Methods</p> <p>-Experimental Methods</p> <p>-Social Network Analysis</p> <p>-Log Analysis</p> <p>-Data Mining and Machine Learning</p> <p>-Mixed methods: convergent design sequential design exploratory design explanatory design</p> <p>-Digital and Computational Methods</p> <p>-Text Mining and Natural Language Processing (NLP)</p> <p>-Network Visualization Tools</p> <p>-Web Analytics</p> <p>-Participatory and Action-Oriented Methods</p> <p>Participatory design</p> <p>Action research</p> <p>-Theoretical and Conceptual Methods</p> <p>-Theoretical modeling</p> <p>-critical analysis</p>	<p>Qualitative methods</p> <p>quantitative methods</p> <p>Mixed methods</p> <p>specialized methods (Usability and user experience studies, Collection assessment, Historical method)</p> <p>Data collection techniques (Direct observation, Indirect observation (e.g., document analysis), Focus group interviews, Archival records analysis)</p> <p>Emerging Approaches:</p> <p>-Knowledge synthesis</p> <p>-Theoretical approaches</p> <p>-Management and operational research</p>	<p>-Quantitative Methods</p> <p>-Surveys</p> <p>-Experiments</p> <p>-Statistical Analysis</p> <p>-Bibliometric Analysis</p> <p>-Qualitative methods:</p> <p>-Interviews</p> <p>-focus groups</p> <p>-cases studies</p> <p>-ethnography -- content analysis</p> <p>-discourse analysis</p> <p>Mixed Methods:</p> <p>-Combining quantitative and qualitative methods to gain a more comprehensive understanding of a research problem.</p>	<p>Quantitative Methods:</p> <p>Experiments</p> <p>Surveys</p> <p>Content Analysis:</p> <p>Statistical Analysis</p> <p>Qualitative methods:</p> <p>-Case Studies</p> <p>-Interviews</p> <p>-Focus Groups</p> <p>Ethnography</p> <p>Mixed Methods</p> <p>Combining Quantitative and Qualitative Methods: Using both numerical data and textual or observational data to achieve a more comprehensive understanding.</p> <p>Other Methods</p> <p>Action Research</p> <p>Design Science Research</p> <p>Discourse Analysis</p>	<p>Quantitative Research Methods:</p> <p>-Bibliometric Analysis</p> <p>-Survey research</p> <p>-experimental studies</p> <p>-Qualitative Research Methods:</p> <p>-Ethnographic Research</p> <p>-Case Study</p> <p>-</p> <p>Phenomenological Studies</p> <p>-Mixed Methods: Convergent Design (combining qualitative and quantitative methods)</p> <p>-Sequential Explanatory Strategies</p> <p>Computational Research Methods:</p> <p>-Network Analysis</p> <p>-Machine Learning</p> <p>-Digital Research Methods: web analytics</p> <p>-Social Media Analysis</p>

Conclusion and key observations

In this study, we used five prompts to assess how five large language models conceptualize and represent information science. As the analysis and comparison above showed, the LLMs in this study provided different responses and representations for the domain of information science with regard to its definition, identity, disciplines, theories, and methods. It should be noted that these LLMs differ in the use of a variety of different datasets, information sources, and machine learning and training models. Therefore, it is not surprising to find that they provide different and inconsistent outputs in response to the prompts related to information science. While they provided accurate conceptions of the interdisciplinary nature of information science, their understanding of the core and widely known theories, methods, and subject areas was inadequate. There are challenges and limitations in how reliably LLMs can be used for domain analysis and disciplinary identification purposes. This study showed that while some LLMs provide more comprehensive and accurate responses and representations, they lack breadth and sometimes relevance in terms of their outputs. These LLMs show deficiencies in their use of credible scholarly and academic sources that could provide a more accurate conceptualization of information science, its definitions, and its theories. Although some resources such as books, Wikipedia, and academic institution websites were used by these LLMs, they were found to be used in random and inconsistent manner. There are also temporal limitations in the coverage of data used to train the LLMs, some are limited to a certain time period, while others lack the historical context, trends, and developments. One key implication and suggestion for the use of LLMs for disciplinary and domain analysis is to utilize multiple LLMs and develop a critical assessment framework for underlying data and output credibility and reliability. Another implication is that the increasing use of large language models in tandem with search engines poses the question of how well LLMs represent knowledge in a domain or more specifically, about concepts and theories within a domain.

References

- Bates, M. J. (1999). The invisible substrate of information science. *Journal of the American society for information science*, 50(12), 1043-1050.
- Belkin, N. J. (1978). Information concepts for information science. *Journal of documentation*, 34(1), 55-85.
- Borko, H. (1968). Information science: what is it? *American documentation*, 19(1), 3-5.
- Boyko, J., Cohen, J., Fox, N., Veiga, M. H., Li, J. I., Liu, J., ... & Xie, X. (2023). An interdisciplinary outlook on large language models for scientific research. *arXiv preprint arXiv:2311.04929*.
- Chua, A. Y., & Yang, C. C. (2008). The shift towards multi-disciplinarity in information science. *Journal of the American Society for Information Science and Technology*, 59(13), 2156-2170.
- Franc, J. M., Cheng, L., Hart, A., Hata, R., & Hertelendy, A. (2024). Repeatability, reproducibility, and diagnostic accuracy of a commercial large language model (ChatGPT) to perform emergency department triage using the Canadian triage and acuity scale. *Canadian*

Journal of Emergency Medicine, 26(1), 40-46.

Hjørland, B. (2002). Domain analysis in information science: eleven approaches—traditional as well as innovative. *Journal of documentation*, 58(4), 422-462.

Janssens, F., Glänzel, W., & De Moor, B. (2008). A hybrid mapping of information science. *Scientometrics*, 75(3), 607-631.

Larivière, V., Sugimoto, C. R., & Cronin, B. (2012). A bibliometric chronicling of library and information science's first hundred years. *Journal of the American Society for Information Science and Technology*, 63(5), 997-1016.

Liang, W., Zhang, Y., Wu, Z., Lepp, H., Ji, W., Zhao, X., ... & Zou, J. Y. (2024a). Mapping the increasing use of llms in scientific papers. *arXiv preprint arXiv:2404.01268*.

Liang, W., Izzo, Z., Zhang, Y., Lepp, H., Cao, H., Zhao, X., ... & Zou, J. Y. (2024b). Monitoring ai-modified content at scale: A case study on the impact of chatgpt on ai conference peer reviews. *arXiv preprint arXiv:2403.07183*.

Liu, Y., He, H., Han, T., Zhang, X., Liu, M., Tian, J., ... & Ge, B. (2024). Understanding llms: A comprehensive overview from training to inference. *arXiv preprint arXiv:2401.02038*.

Mitchell, M., & Krakauer, D. C. (2023). The debate over understanding in AI's large language models. *Proceedings of the National Academy of Sciences*, 120(13), e2215907120.

Petras, V. (2024). The identity of information science. *Journal of Documentation*, 80(3).

Wang, J., Hu, H., Wang, Z., Yan, S., Sheng, Y., & He, D. (2024, May). Evaluating Large Language Models on Academic Literature Understanding and Review: An Empirical Study among Early-stage Scholars. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-18).

Wijewickrema, M. (2023). A bibliometric study on library and information science and information systems literature during 2010–2019. *Library Hi Tech*, 41(2), 595-621.

Wu, K., Wu, E., Cassasola, A., Zhang, A., Wei, K., Nguyen, T., ... & Zou, J. (2024). How well do LLMs cite relevant medical references? An evaluation framework and analyses. *arXiv preprint arXiv:2402.02008*.

Xu, H., Guo, T., Yue, Z., Ru, L., & Fang, S. (2016). Interdisciplinary topics of information science: a study based on the terms interdisciplinarity index series. *Scientometrics*, 106, 583-601.

y Arcas, B. A. (2022). Do large language models understand us?. *Daedalus*, 151(2), 183-197.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., ... & Wen, J. R. (2023). A survey of large language models. *arXiv preprint arXiv:2303.18223*.

Zins, C. (2007). Knowledge map of information science. *Journal of the American Society for Information Science and Technology*, 58(4), 526-535.