

# Representationalism in Knowledge Organization and Information Retrieval: Cheer or Criticism?

Mahboobeh Farashbashi Astaneh

Ferdowsi University of Mashhad, Information Center & Central Library,  
Ferdowsi University of Mashhad, Azadi Square – Mashhad – Iran, <astaneh@um.ac.ir>

Mahboobeh Farashbashi Astaneh holds a PhD in Knowledge & Information Studies from Ferdowsi University of Mashhad, Iran. Her research interests encompass knowledge organization, semantics, philosophy of language, and information retrieval. Presently, alongside her position at the university's central library, she also serves as a teaching faculty member in the Faculty of Education & Psychology at Ferdowsi University of Mashhad.



Astaneh, Mahboobeh Farashbashi. 2023. "Representationalism in Knowledge Organization and Information Retrieval: Cheer Or Criticism?". *Knowledge Organization* 50 (7): 475-488. 34 references. DOI:10.5771/0943-7444-2023-7-475.

**Abstract:** The purpose of this paper is to critically analyze the representational approach to language using the concept of 'Representationalism' in philosophy and a new framework referred to in this paper as Information Retrieval Action (IRA). IRA is a theoretical construct based on Information Retrieval (IR) theory. Therefore, the methodology employed in this study is philosophical and speculative. It examines and discusses representationalism in both the mental (cognitive) and linguistic dimensions, incorporating concepts from the philosophy of mind and the philosophy of language as methodological approaches. 'Concept' and 'Meaning' emerge as two critical semantic elements when considering language as a semiotic phenomenon. Representationalism in IRA is rooted in logical positivism, which serves as the foundation for interpreting meaning and concept based on mental representation. In response, the paper proposes an alternative non-representational approach to the semantic elements of language (meaning and concept) based on the philosophy of pragmatism within the IRA framework. While the emergence of new technologies like ontologies implicitly criticizes representationalism in IRA, previous research has not explicitly addressed the theoretical and philosophical criticism of representationalism within the IRA framework.

Received: 14 March 2023; Revised 07 October 2023; Accepted 13 October 2023

Keywords: representationalism; information retrieval theory; knowledge representation; pragmatism; language

## 1. Introduction

Information retrieval (IR) has been studied as one of the main fields of information science from various aspects. For example, it has been approached from technical and mathematical perspectives in computer science. In cognitive sciences, particularly psychology, the focus has been on human and cognitive aspects. In knowledge and information science, the emphasis has been on indexing and organization. Consequently, IR research has been associated with various names, such as IR systems, IR interactions, IR behavior, and IR theory. However, IR theory has generally been considered a technical theory related to computer science per se. It has not been studied from a theoretical perspective, prompting Kelly (2019) to call for more theory to be developed in this area. This leads to a vague image of IR in Library and Information Science (LIS) field, as Thornley and

Gibb (2007) inquire whether information retrieval (IR), as a paradoxical concept, can be better understood through dialectical analysis, addressing the role of meaning. According to them, to comprehend IR's problematic nature, it is necessary to delve into its core concept and meaning as a dialectical process. They contend that philosophy can unveil the conflicts and contradictions in IR.

Hjørland (2021) has emphasized the importance of IR from a philosophical standpoint. He states that "Information retrieval (IR) and knowledge organization (KO) are two research fields that, on the one hand, are separate fields of study, but on the other hand, share the same aim: to facilitate the findability of documents, knowledge, and information" (1). He argues that we need to base IR on KO and derive insights from the philosophy of science to develop a theoretical understanding of IR. He previously examined the "relation between KO and Semantics" and considered

the “tools developed in KO for IR as basically semantic tools” (Hjørland 2007, 367). In this context, Mai (2001) quotes Blair, stating that “the central task of IR... is to understand how documents should be represented for effective retrieval. This is primarily a problem of language and meaning. Any theory of document representation... must be based on a clear theory of language and meaning” (592). According to Mai, the fundamental form of KO revolves around the connection of language with meaning. Thornley and Gibb (2007, 761) claim that these two aspects are fundamentally different:

[...] the individual experience of meaning can both be described in language and is dependent on a shared objective language. The subjective and objective aspects of meaning are therefore both fundamentally different in some ways, as the divided research tradition in IR testifies, but also in a relation of uneasy dependence.

The authors (2007) argue that a dialectical understanding of meaning can help elucidate certain aspects of the complex nature of IR. They attribute these contradictions to the dialectical relationships inherent in meaning. In their view, the dialectical relationship is characterized by “mutual dependence and antagonism, where an understanding of each opposing aspect necessitates an understanding of how it relies on its negation” (756). Therefore, employing the philosophy of meaning requires a comprehensive grasp of the philosophical oppositions underlying the nature of meaning. Referring to Putnam’s view, they point out that “meaning has two aspects: objective and subjective. Within the philosophy of language, the subjective aspect of meaning is often referred to as intension, and the objective aspect of meaning as extension” (758). One may ask why we need to investigate meaning representation from the IR perspective when much research has already been conducted in the KO area. This is due to the inherent interconnection between IR and KO in representing knowledge embedded within language. While KO plays a significant role in structuring and categorizing information, IR provides a complementary and essential dimension for a comprehensive understanding of meaning.

Firstly, IR bridges the gap between the abstract nature of knowledge representation and its practical application. We can effectively retrieve, and access information based on its meaning, relevance, and context by employing IR techniques. This process involves understanding the user’s information needs and utilizing retrieval algorithms to match those needs with the most suitable and meaningful information. Secondly, investigating meaning necessitates a consideration of the dynamic and context-dependent nature of language use. IR allows us to explore the intricate relation-

ship between language, context, and social and cultural factors. By analyzing the patterns and associations within a corpus of information, we can uncover more profound layers of meaning beyond the static categorization provided by KO alone. Furthermore, IR complements KO by providing a user-centric approach. While KO focuses on organizing knowledge based on predefined categories and structures, IR acknowledges the individual’s subjective information needs and preferences.

Overall, investigating meaning needs the integration of both IR and KO. While KO provides a foundation for structuring information, IR enables us to actively retrieve, analyze, and interpret that information to capture the dynamic nature of meaning. By combining these approaches, we can better understand language representation elements (meaning and concept) that align with the philosophical aspects of language, context, and user-centered exploration.

This study explores the understanding and analysis of representationalism within a new framework, aiming to critique existing approaches and suggest a new language-based approach for representing knowledge in LIS. Since its objective is to examine the representational aspects of the semantic elements of language, namely meaning and concept<sup>[1]</sup> in IR theory, it is essential to approach IR theory from a philosophical perspective, precisely the philosophy of language within the framework referred to as Information Retrieval Action (IRA), hereafter referred to as IRA.

The IRA framework is a conceptual model that, on the one hand, offers a structure for arranging and examining the procedures associated with IR, and on the other hand, it can demonstrate the representational aspect of language elements (meaning and concept) throughout the entire process of KO and IR. Even though the IRA framework concentrates on a theoretical structure aiming to merge various theoretical viewpoints that underlie assumptions and philosophical consequences of IR established in KO, it could serve as a temporary alternative to IR theory. Specifically, the term “Action” encompasses practical activities like indexing, organizing, and retrieving information and cognitive and emotional processes such as user intention, assessment of relevance, and feedback. Since the theoretical aspects of KO and IR constitute IRA, they are essential keys to understanding representationalism in LIS, as shown in Figure 1.

The relationship between IRA and representationalism lies in their connection to the representation of knowledge within the realm of language encompassing meaning and concept elements. Technically speaking, the theoretical aspects of IRA for understanding representationalism require delving into the role of language and its semantic elements (meaning and concept). Few studies have addressed this aspect of IR theory (Blair 1990, Blair 1992, Blair 2003, Thornley and Gibb 2007, Thornley and Gibb 2009).



Figure 1. An illustration of IRA for representationalism in LIS

Therefore, this paper intends to explore and analyze representationalism in the semantic layer of language, including meaning and concept, within IRA. This will be done using a theoretical approach and employing both philosophical and speculative approaches. The philosophical method involves logical reasoning, conceptual analysis, argumentation, and examining evidence and premises to arrive at rational conclusions. It often relies on logical arguments, thought experiments, and the examination of various perspectives to gain insights into abstract and complex subjects. A speculative approach refers to a method of inquiry that involves engaging in speculative or theoretical reasoning to explore abstract concepts or the nature of reality. It often involves conceptual frameworks, thought experiments, and hypothetical scenarios to gain insights into the nature of the world, the mind, or other philosophical domains. Speculative approaches are characterized by their emphasis on abstract reasoning and the exploration of possibilities and potentialities.

The discussion begins by exploring the roots of representation in the semantic layer of language, which is traced back to representationalism in the philosophy of language. The concept of representationalism is discussed, and studies are reconceptualized based on their concern for language and meaning. The representational approaches to language in IRA are analyzed in three parts, and at the end, an alternative approach to meaning and concept representation in IRA is presented.

## 2.0 Representationalism

The concepts of representamen, representation, and representationalism are generally referred to as “representational-

ism” or “cognitive representation theory,” “perceptual representation theory,” or “indirect realism.” From an in-depth perspective, representationalism is generally rooted in cognitive sciences and the philosophy of mind in particular. The philosophical theory of cognition is based on the claim that the mind cannot perceive objects directly from the outside but only through mental forms, perceptions, or representations of material objects, not the material objects outside the mind itself. According to Tye (2002, 137), representationalism is a concept that deals with the subjective “feel” or phenomenal character of experiences. At its core, the thesis of representationalism is one of supervenience, which means that experiences with similar representational contents also share similar phenomenal character. However, this thesis does not provide any information about the nature of phenomenal character itself. Strong or pure representationalism takes things a step further by attempting to define what a phenomenal character is. Tye’s theory, developed in 1995, posits that phenomenal character is identical to representational content that satisfies certain additional conditions.

Egan (2012) outlines various forms of representationalism within cognitive science and describes the most widely accepted form as the belief that the human mind is an information-processing system and that human cognitive abilities can be understood as representational capacities. Specifically, representationalism holds that the mind uses mental representations to process information. Representationalism generally holds that humans perceive the external world as an object that exists independently of the mind. However, the mind does not perceive the external world directly but rather constructs mental representations of sense-data through perception. This is why representationalism is also

known as indirect realism in the philosophy of mind. While this is the basic premise of representationalism in cognitive science and philosophy of mind, this paper will focus on the representational approach to language (meaning and concept) and will critique this approach in the context of IRA.

Looking back at the origins of representationalism and the representational approach to language, we can trace it back to seventeenth-century philosophers such as Descartes, Locke, and Leibniz. During this period, the focus was primarily on the mind rather than language. According to Hirst (2006, 265), the classic form of this view was the “representative realism” espoused by Descartes and Locke, which is still upheld by many scientists today. This view posits that the mind constructs mental representations of the external world based on sensory input, rather than directly perceiving the external world itself. According to Yang (2008), early modern philosophy placed a greater emphasis on epistemology rather than language, focusing on *Cogito*. As a result, language was not given clear and specific consideration. The prevailing attitude toward language during this period was that it acted as an intermediary connecting subject with object and expressed the meanings of the object, but it was not necessary to enhance its status. Since early modern philosophy did not concern itself with intersubjectivity or the relationship between “I” and “the other,” there was no need to consider the function of language as a bridge between subjects or consciousnesses, let alone among cultures (596-97).

Representationalism in early modern philosophy can be described as “representational realism” or “indirect realism” of language. According to this view, the mind perceives things through the senses from the real world, and language reflects these mental impressions equally, which is why this theory is often referred to as indirect realism. However, in the twentieth century, both the analytic and continental traditions of philosophy focused on linguistic meaning and how language relates to reality (Crane 2005). Indeed, the shift towards focusing on the role of language in shaping our understanding of the world occurred in analytic philos-

ophy, which drew inspiration from the groundbreaking work of Frege<sup>[2]</sup> and Russell<sup>[3]</sup> in logic. In this view, a word’s meaning is related to things in the world and to other words with which it combines to make sentences. This dual relationship can be seen in Figure 2, which illustrates how words are related to both the external world and other words in sentences.

From a philosophical perspective, representationalism emphasizes that knowledge is represented through symbols or mental representations. It suggests that these representations mediate our understanding of the world. This perspective allows a deeper understanding of the complex relationship between language and reality. By considering the dual relationship of words to both the external world and other words in sentences, this perspective demonstrates how words serve as a bridge between our understanding of the external world and our ability to express that understanding through language. It also highlights the importance of context in determining the meaning of words, as the meaning of a word can vary depending on the words with which it is combined in a sentence. Overall, this approach paves the way for a more nuanced understanding of the role of language in shaping our perception and understanding of reality. In this view, the semantic elements of language (meaning and concept) are considered to be dependent on the mind, making this philosophical approach relatively static.

As discussed earlier, representationalism in philosophy generally associates language with the mind, regarding it as a means through which mental representations of the external world are constructed. It underscores the significance of mental representations in our understanding and interaction with the world, suggesting that mental representations of objects, events, and concepts mediate our experience and knowledge. Representationalism offers a theoretical foundation and framework for comprehending how knowledge is represented within the language domain, encompassing both meaning and concept. It facilitates understanding how language elements (meaning and concept) are represented

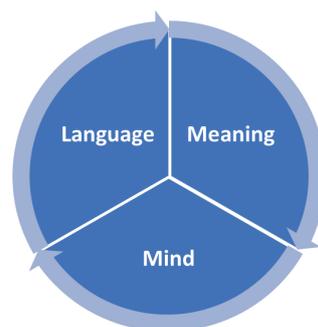


Figure 2. Understanding representationalism in philosophy

and interpreted within the IRA framework. Furthermore, it has the potential to significantly contribute to the comprehension and advancement of the theory of IRA, particularly in the context of knowledge representation. As the paper aims to analyze the representational approach to the semantic elements of language, representationalism in IRA will be discussed in the next section.

### 3.0 Representationalism in Information Retrieval Action (IRA)

#### 3.1 Definition

Representationalism is not a specific concept or methodology developed exclusively for IRA. Rather, it is a broader philosophical perspective applied to various domains, including IR, to enhance the understanding and interaction with information. In the context of IRA, the representational approach to language and meaning involves using language as a medium to represent knowledge in a format suitable for processing by an IR system.

Ingwersen's distinction between the two fundamental forms of representation in information science - 'formal' and 'topical' (1992, 53) - provides a useful framework for understanding how information is organized and retrieved. Formal representation, also known as bibliographic representation, focuses on the metadata associated with a document, such as its title, author, date of publication, and other descriptive information. This type of representation is particularly helpful for identifying and retrieving documents based on their bibliographic characteristics, such as the author or publication date. On the other hand, topical representation is concerned with the content of the document and its subject matter. This type of representation is particularly useful for identifying and retrieving documents based on their content, such as keywords, concepts, or themes. Both formal and topical representation play important roles in information science and are often used in combination to facilitate effective IR. For example, a search engine may utilize both the formal metadata and the topical content of a document to determine its relevance to a user's query.

Chu's (2010) contribution to understanding information representation is highly valuable, particularly regarding information organization and retrieval. In building upon Ingwersen's framework, Chu identifies several key methods of information representation. These methods include abstracting, indexing, categorization, summarization, and extraction. Abstracting involves generating a concise summary or abstract of a document, capturing its essence through keywords or phrases. On the other hand, indexing entails assigning descriptors or subject headings to facilitate categorization and improve retrieval. Categorization involves organizing documents into specific categories or sub-

ject areas based on their content. Summarization focuses on condensing a document to highlight its key points effectively. Additionally, extraction plays a crucial role by automatically identifying and extracting relevant information from a document, such as named entities or significant phrases. By employing a combination of these representation methods, information can be structured in a way that simplifies organization, search, and retrieval processes. The effectiveness of information representation is vital for the success of various information management systems, including search engines, digital libraries, and content management systems. Chu's insights contribute to enhancing these systems' functionality and usability by optimizing information representation techniques.

Friedman and Thellefsen's (2011) observations regarding the notion of representation and knowledge representation in KO research are pertinent to the ongoing debates and challenges in the field. They highlight that the term 'representation' carries a strong inherent realism, suggesting that the object being represented possesses an independent existence outside of its representation. This can pose difficulties for KO researchers who aim to comprehend how knowledge is represented and organized in diverse contexts. As a result, some researchers have proposed alternative terms, such as 'depiction,' 'manifestation,' or 'expression,' to avoid the connotations of realism associated with representation. Furthermore, Friedman and Thellefsen argue that the absence of a unified terminology and theoretical understanding of knowledge representation has contributed to the challenges faced by KO researchers. They suggest that a more structured and theoretical approach to knowledge representation is necessary for advancing the field. This entails developing a common vocabulary and conceptual framework for comprehending knowledge representation, as well as exploring the relationships between different types of representations, such as linguistic, visual, and symbolic representations. Overall, Friedman and Thellefsen's work emphasizes the significance of critically examining the assumptions and concepts underlying knowledge representation in KO research, and cultivating a more nuanced and theoretically grounded understanding of this complex and multifaceted phenomenon.

Thornley and Gibb's (2009) description of document representation underscores IR's challenges, particularly the potential for failures in recall and precision. When a document representation and a query share subject matter but use different terms to describe it, there is a risk of recall failure, meaning that the document may not be retrieved even though it is relevant to the user's information need. This can occur due to variability in the terminology used to describe a specific subject or due to differences in perspectives between the indexer and the user regarding what constitutes relevant information. Conversely, when a document repre-

sensation and a query use the same terms but have different meanings due to their respective contexts, there is a risk of precision failure, meaning that the retrieved document may not be relevant to the user's information need. This can happen when the same term is employed to describe different concepts in different contexts or when the user's information need is more specific than the information in the retrieved document. These challenges underscore the importance of developing effective document representation and query formulation techniques. Furthermore, techniques for matching queries to documents must account for different perspectives and contexts. This can involve applying natural language processing techniques, machine learning algorithms, and other computational approaches that help identify and resolve differences in terminology and context (2009).

Mai (2000, 94) describes the process of representing knowledge or documents, which involves "capturing the subject matter, content, topic, or aboutness of the documents". This process is crucial for document retrieval as it assists users in finding relevant documents by matching their search queries with the subject representation of the documents. The author also distinguishes between two types of representation: descriptive representation and subject representation. Descriptive representation focuses on describing the physical characteristics of a document, such as its title, author, publisher, and date of publication. It also identifies the responsibility for the intellectual content of the document, such as the author or editor. On the other hand, subject representation concentrates on identifying the subject concepts covered by the intellectual content of a document. This involves analyzing the text of the document to determine its main topics, themes, and ideas, and representing these concepts in a way that facilitates document retrieval. Subject representation is often achieved through the use of controlled vocabularies, such as thesauri or subject headings, which provide standardized terms for describing the content of documents. Still, according to Mai (2000), it is important to differentiate between the descriptive and analytic aspects of representation, although this study focuses on the latter. Specifically, the definitions of representation discussed here—such as thematic, conceptual, and meaning representation—address how knowledge is represented rather than the physical representation of documents or information. Language plays a crucial role in organizing and accessing information, which is why studying the representation of knowledge from an epistemological perspective is significant. This approach is known as representationalism, which is based on the idea that language serves as a tool for representing knowledge. The next three sections will delve further into this representational approach to language.

As can be seen, these research studies provide a comprehensive overview of various aspects related to knowledge rep-

resentation. They distinguish between two key aspects of representation: bibliographic representation and subject representation. Bibliographic representation focuses on the descriptive details of a document, whereas subject representation delves into the identification and representation of the subject concepts covered by the intellectual content of the document. The latter involves analyzing the text to determine the main topics, themes, and ideas and representing them to facilitate document retrieval. Controlled vocabularies, such as thesauri or subject headings, are often used to standardize the terms used to describe document content. The studies also acknowledge the significance of differentiating between descriptive and analytic aspects of representation, with a particular focus on the latter. It explores various definitions of representation, including thematic, conceptual, and meaning representation, which primarily address how knowledge is represented rather than the physical representation of documents or information. Language is recognized as a crucial component in organizing and accessing information, emphasizing the relevance of studying knowledge representation from an epistemological perspective.

Overall, these studies provide a solid foundation for the subsequent sections of the paper, which will delve further into the representational approach to language. It paves the way for a comprehensive investigation of representationalism and the implications for understanding how language serves as a tool for representing knowledge.

### 3.2 Understanding

When it comes to understanding representationalism in IRA, it is essential to begin with language. Language plays a crucial role in this discussion because any interpretation of language can impact the meaning and concept associated with it. In fact, language forms the most significant component in representing knowledge within IR tools since knowledge is presented through concepts, and these concepts create meaning.

Mai (2001) highlights the roots of representationalism in IRA. He argues that "indexing is not a neutral and objective representation of a document, but rather a representation of an interpretation of the document for future use" (591). According to Mai, indexing concerns meaning and language and involves interpreting and representing documents. He emphasizes that different words and phrases are produced differently in each person's context. To analyze the stages of the indexing and classification process more comprehensively, Mai uses Peirce's theory of signs and highlights Blair's view that Wittgenstein's philosophy of language is highly relevant to understanding these processes. However, Blair (1990) rejects semiotics as a foundation for understanding indexing and IR. Instead, he argues that Wittgenstein's later theories provide a valuable framework for representing doc-

uments for retrieval. Mai notes that while Blair doesn't endorse Peirce's semiotics, there are similarities between Peirce's semiotics and Wittgenstein's pragmatic philosophy of language.

Weiss et al. (2016) have defined representationalism in IRA as the belief that a stable relationship exists between words and the world they represent. In other words, words can accurately represent aspects of the world. Day (2016) has interpreted this representational approach to language by adapting Paul Otlet's information organization methods to contemporary IRA. Otlet's epistemology is commonly referred to as a 'picture theory' of knowledge in contemporary philosophy, "which holds that truth lies in the correspondence of statements with states of the world" (2016, 3). Day notes that this type of representationalism or 'positivism' is distinct from the naïve empiricism of British philosophers and experimentalists like John Locke. He sees representationalism as comparable to positivism and equates it with the picture theory of knowledge. Similarly, Hjørland (1998) also connected IR problems to semantic theories and explored the relationship between the picture theory of meaning and theoretical assumptions in IR in detail.

Svenonius (2004) also linked the representation of knowledge in IR languages to theories of meaning and emphasized the connection between representation and meaning. She identified the picture theory of meaning as a representational theory of meaning and situated the representational approach to meaning within the epistemological foundations of logical positivism.

Thornley and Gibb (2009, 135) suggest that in some cases, "understanding meaning requires examining its reference or content, but this process can only be grasped within the broader context of its use and purpose". In their 2009 publication, the authors delve into the evolution of the representation of meaning in philosophy, starting from Frege up to Wittgenstein. Frege's approach to understanding meaning involved identifying qualities within objects and our mental experiences of understanding words that hold meaning. Wittgenstein, on the other hand, proposed that we should pay attention to how meaning is used, rather than the relationship between different entities.

Based on studies and the views of language philosophers, it appears that representationalism in IRA is rooted in logical positivism, which dominated the field for many years. The "picture theory" of Wittgenstein is one manifestation of this positivist paradigm, as reflected in various studies in the field by researchers such as Svenonius (2004) and Blair (2003). As Reese (1996, 393) notes, "in the empiricist movement, the analysis of language was equated with moving from impressions to ideas and the use of signs in reality". Therefore, to understand representationalism in IRA, one must consider the assumptions of positivism in practice

### 3.3 Analysis

To understand the assumptions of positivism in the analysis of representationalism, it is important to examine language as a semantic system. This approach emphasizes the central role of "meaning" and "concept" in language, which is crucial for analyzing representationalism within the context of IRA. Studies by Rafferty (2001) and Huang (2006) compare meaning to semantic and semiotic issues. Their research links conceptual and semantic representation issues to semiotics. For instance, Rafferty (2001) investigated the semiotic role of language in classification systems and noted that classification notation is a semiotic activity based on language as a system of signs. The subject term, which is the signified, is itself a symbolic signifier at the level of language. The classification system is further complicated by the fact that it is based on the analysis and description of knowledge recorded in documents, which are interpreted by human indexers. Therefore, "the issue of interpretation and matching plays a significant role in IR" (187). Rafferty's study implicitly uses the semiotic role of language to analyze knowledge representation in IRA.

Huang (2006) explores the relationship between representation, semiotics, and concept theory and argues that information science and semiotics are both concerned with how humans connect representation and meaning. He refers to the semiotic triangle, which illustrates the relationships between the user of information, the representation of information, and the meaning of information. According to Huang, information is created by users and acts as a link between understanding symbols or represented objects. Drawing on Raber and Budd's perspective, Huang suggests that two fundamental problems in IR are

[...] to assign accurate and adequate representative descriptions for an informative object upfront, and the other is to assess the relevance of retrieved results on the backend. These two problems are akin to the concern with language that is used as a tool of systematic inquiry for knowledge discussed in semiotics (8).

Friedman and Thellefsen (2011) investigated the concept of knowledge representation by comparing Peirce's semiotic theory and Dahlberg's concept theory. While both theories are concerned with knowledge representation, they differ in their approaches. Peirce's semiotic theory focuses on the philosophical and logical aspects of how signs create meaning in the mind of a perceiver. It offers tools for interpreting and understanding the representation process but is not specific to any particular knowledge representation system.

In contrast, Dahlberg's concept theory examines the relationships between objects and their classification, providing a method for analyzing and organizing concepts within

a knowledge organization system (KOS). Although semiotics offers a philosophical context for the concept of representation and distinguishes between the immediate and dynamical objects, Dahlberg's theory provides a practical method for analyzing representation and concepts in a KOS environment.

To summarize, Friedman and Thellefsen (2011) compare Peirce's semiotic theory and Dahlberg's concept theory in the context of knowledge representation. Peirce's semiotics is focused on how signs can convey meaning to a perceiving mind, while Dahlberg's theory examines concepts and their representation by considering the relationship between the concept and its classification. Dahlberg's theory provides a method for analyzing representation and concepts within a KOS environment, while semiotics offers a philosophical context for the concept of representation. Additionally, semiotics distinguishes between the immediate object (the sign) and the dynamical object (the object being represented), providing a clear understanding of the function of a representation of an object and the object itself. Friedman and Thellefsen suggest that "Peirce's semiotic theory and Dahlberg's concept theory offer distinct perspectives on knowledge representation, particularly in relation to concepts and entities in KO" (645). They emphasize the significance of these theories for understanding knowledge representation within the broader context of KO and highlight the value of combining the concepts of "knowledge representation," "semiotic theory," and "concept theory" in the context of KO. Hence, they argue that a deeper understanding of knowledge representation can be achieved.

Figure 3 illustrates the analysis of representationalism within IRA framework, highlighting its key components. When language is regarded as a communication system, both the elements of "meaning" and "concept" find their place within the semantic layer. According to a representational approach to "meaning" and "concept" in IRA, the mental and cognitive aspects of these elements are crucial for understanding representationalism within the context of IRA. This implies that the approach to the semantic elements of language is firmly grounded in positivism in the philosophy of language, establishing a profound interconnection with cognitive science. This underscores the criticality of mental representation of the concepts. The figure may also suggest that positivism and cognitivism, although rooted in empiricism, are different traditions with the same formal logic. Smythe (1992) argues that positivism and cognitivism are distinct traditions sharing a common formal logic rooted in empiricism, which reflects the external world. Overall, the figure outlines a framework for understanding representationalism in IRA, incorporating positivism and cognitive science perspectives and emphasizing the mental and cognitive aspects of meaning and concept.

### 3.4 Challenges

According to Smythe (1992), the cognitivist theory of mind is rooted in the idea of interpreted formal systems. This theory posits that the mind functions as a computational system that interprets and utilizes formal symbols and rules, emphasizing the study of how individuals process and manipulate information based on the interpretation of these symbols and rules. Smythe's argument underscores the significance of interpreted formal systems in the cognitivist theory of mind, viewing the mind as a computational system that employs formal rules and symbols to process information. This perspective has greatly influenced the fields of artificial intelligence and cognitive science, providing a framework for understanding how machines and humans process information in similar ways.

However, as Frohmann's critique suggests, this perspective has limitations in that it tends to overlook the social and cultural factors that influence the interpretation and use of formal symbols and rules. By considering the social and historical context in which information processing occurs, we can achieve a more comprehensive understanding of how individuals and communities utilize formal systems to organize and comprehend information.

Three decades ago, Frohmann (1990) criticized mentalism in IR theory, issuing a warning about representationalism, detailed indexing, and their implicit rules:

According to mentalism, rules can have no justification, because they are simply given pieces of our mental equipment, whether they are innate and universal, or effects and particular. Rules of indexing are rules of text representation for the purpose of text retrieval. But text retrieval designates a set of particular social practices. Consequently, the construction of indexing rules institutes or facilitates particular kinds of retrieval practices and depends, therefore, upon a preliminary understanding of the social practices constituting text retrieval in the actual, historically real social world (97).

Frohmann's argument against mentalism in IR theory is rooted in the idea that the rules of indexing in text retrieval systems are socially constructed and variable, rather than universal and innate. He suggests that these rules are not simply given pieces of mental equipment but are shaped by the specific social practices and needs of the community using the retrieval system. Frohmann's critique challenges the notion that mental states and processes are the primary objects of study in understanding human behavior. Instead, he argues that social and historical context plays a major role in shaping the way individuals and communities organize and retrieve information. This perspective emphasizes the im-

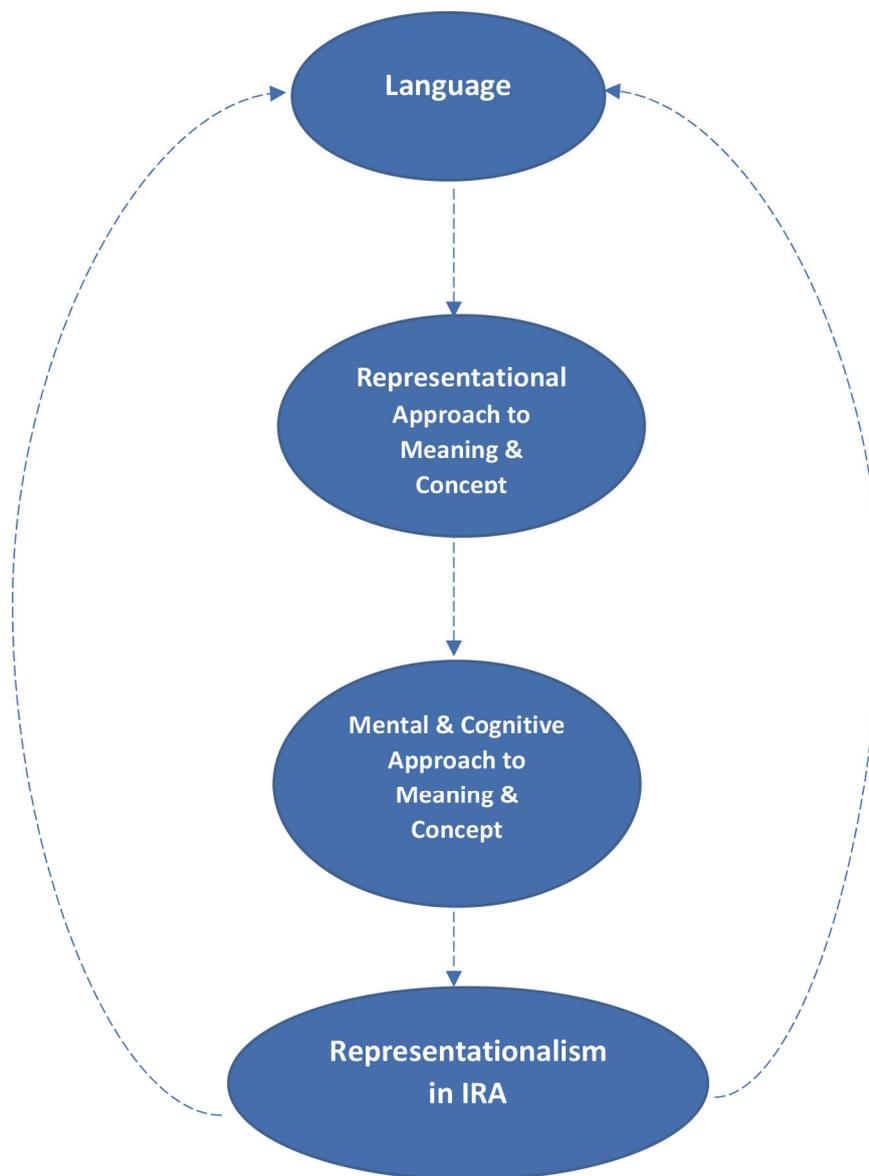


Figure 3: Analysis of Representationalism in IRA

portance of studying the social and cultural factors that influence the development and use of information systems, rather than focusing solely on individual cognitive processes. Overall, Frohmann’s critique highlights the need for a more nuanced and socially aware approach to understanding the nature of information and knowledge organization. By recognizing the social and historical context in which information systems are created and used, researchers can better understand the complex interactions between individuals, communities, and the information they rely on.

Ellis (1992) argued that the failure of theory development in IR is due to the “inherent categorial duality” (60) between the physical and cognitive paradigms in IR research. On one hand, IR deals with the physical paradigm, which involves inherently inflexible features of technical is-

ues. On the other hand, IR also deals with the mental and cognitive world of users and the human and interactive aspects of IR. According to Ellis, both the physical and cognitive approaches in developing and expanding IR theory have been unsuccessful because they fail to integrate these two paradigms fully. Instead, he emphasized the need for strong theoretical approaches to bridge the gap between IR’s physical and cognitive aspects. Ellis’s perspective highlights the importance of developing theoretical frameworks that can account for IR’s technical and human aspects. By considering the complex interactions between technical systems and human users, researchers can develop more effective and comprehensive IR approaches to meet users’ needs better.

Hjørland (2013) criticized the user-centered approach and cognitive perspective that have been influential in infor-

mation science since the 1980s. He points out that this paradigm has been overly focused on individual cognitive processes and has largely ignored the role of culture and society in shaping these processes. According to Hjørland, many scholars became skeptical about the theoretical basis of the cognitive paradigm after 1990, as they began to realize that the role of culture and society in cognition was being marginalized or ignored. He suggests that this approach tends to view IR as a purely technical problem to be solved by optimizing the interaction between users and systems rather than recognizing it as a complex social and cultural phenomenon. Hjørland's critique emphasizes the need for a more socially and culturally aware approach to IR that considers the broader societal and cultural factors influencing how individuals use and interact with information.

Hjørland (2011) criticized a study that compared automatic indexing and human indexing and concluded that human indexing is not superior to automatic indexing. He argues that the study is based on a representational, mind-based, or cognitive perspective that advocates for the cognitive approach, which many researchers, including himself, have criticized. According to Hjørland, cognitive views of indexing suggest that people index and search documents in a specific way because they have a certain cognitive or mental structure that cognitive studies may uncover and provide a basis for indexing. He suggests that this perspective implies that the indexing rules are parts of our cognitive structures, which are connected to universal, biological given structures. However, he argues that cognitive indexing views are theoretically unclear and problematic. Hjørland's critique highlights the limitations of the cognitive approach to IR, which tends to focus on individual cognitive processes and overlook the broader social and cultural factors that shape how we organize and search for information. By recognizing the importance of these factors, researchers can develop more comprehensive and effective approaches to IR that better reflect the complex social and cultural contexts in which information is created, shared, and used.

Blair (2003) has investigated the challenges of representationalism in information retrieval, which seeks to go beyond cognitive and linguistic issues, particularly in the philosophy of language. In his analysis, Blair explores the influence of the philosophy of language on IR and discusses the reasons for the failure of the representational approach to language in describing and differentiating information. Blair delves into the nature of meaning and notes that philosophers have long pondered the "meaning of meaning," with Wittgenstein's work having a particularly influential role in the philosophy of language during the twentieth century. He attributes the emergence of the linguistic turn in analytic philosophy to Wittgenstein's later work, *Philosophical Investigations*. According to Blair, the linguistic turn arose because philosophers who intended to study ideas fo-

cused on describing ideas, not just what we think, but also what we say we think. Blair quotes Hacker and Rorty, stating that the only way to access ideas directly is through our own expressions.

Blair's analysis highlights the limitations of the representational approach to language in IR and emphasizes the need to consider broader social and cultural factors that shape how we describe and differentiate information, going beyond purely cognitive and linguistic issues.

Hjørland (2004) explores the foundations of representational challenges in IR rooted in philosophical problems, particularly those associated with realism. He criticizes the traditions of empiricism and positivism that govern IR, referring to them as antirealism, which contrasts with the realist claim that an independent reality exists apart from the mind. Hjørland argues that many people mistakenly equate empiricism and positivism with realism when these traditions strongly advocate for antirealism. He suggests that a clear distinction between empiricism/positivism and realism should be drawn. Hjørland also implies that the concept of representationalism is closely linked to antirealism and criticizes specific disciplines, such as psychology, linguistics, and artificial intelligence, for their antirealist approaches to cognition. He asserts that this antirealism is based on the representational theory of perception, which focuses on individual cognition while disregarding the broader social, cultural, and historical implications of human cognition. Hjørland's critique underscores the significance of realism in IR and emphasizes the need to consider the broader social and cultural contexts in which information is created, shared, and utilized. By acknowledging the influence of philosophical problems, particularly those related to realism on IR, researchers can develop more comprehensive and effective approaches that better capture the intricate social and cultural landscapes in which information is situated.

In conclusion, although views may differ, they collectively demonstrate that the representation of knowledge in the representational approach to the semantic elements of language is primarily associated with the linguistic paradigm and cognitive science. By acknowledging the impact of the philosophy of language on IRA, researchers can devise more comprehensive and efficient methodologies that better capture the intricate social and cultural landscapes in which information is generated, exchanged, and utilized. Now the question is, how can "meaning" and "concept" be interpreted within the context of IRA while making a transition from representationalism and still being committed to it at the same time? The last part of this discussion will address whether a non-representational approach can be effective for the accurate interpretation of meaning and concept.

#### 4.0 Non-Representational Approaches as Alternatives in IRA

According to the Stanford Encyclopedia of Philosophy, concepts are the building blocks of thoughts, and in knowledge organization systems (KOS), they serve as the foundation for knowledge representation. Similarly, Stock (2010) explains that concepts are determined by their intension, extension, and definition, and they are essential components of language representation systems. However, when analyzing IR using a representational approach to language, the semiotic aspect of language appears to take precedence, and concepts become associated with signs to create meaning. In this approach, meaning is relational and relies on the connection between a sign and a concept established through a code. This representational approach to language can lead to mentalism in cognitive science and positivism in the philosophy of language. Therefore, a non-representational approach, such as a pragmatic view, is necessary within the context of IRA. Non-representational approaches align with the pragmatic view that meaning and concept are shaped by communicative intentions and the context in which communication occurs. They emphasize the importance of considering the speaker’s intended meaning and the listener’s interpretation in light of the communicative context. By adopting a pragmatic understanding of meaning and concept, the non-representational approach can enhance accuracy by considering language use’s dynamic and situated nature. In other words, a pragmatic view considers language in its cultural and social context, recognizing that language is a dynamic phenomenon shaped by social context, culture, and community. It also counts meaning as a pragmatic concept. A pragmatic view of meaning recog-

nizes language as a dynamic, context-dependent, and social phenomenon. It acknowledges that meaning is not solely derived from the words themselves but is constructed through the interaction between speakers, listeners, and the situational context.

In a non-representational approach, “concept” and “meaning” are distinct elements that derive their existence from social context and culture. This approach leads to a pragmatic understanding of concepts rather than a tautological or solely mental inference of the concept. The non-representational approach to language and meaning offers a more nuanced and contextually sensitive understanding of concepts and their relationship to meaning in IR. Figure 4 illustrates representational and non-representational approaches to semantic elements of language, namely “concept” and “meaning” in IRA.

From a technical standpoint, the representational approach to language posits that there is a direct correspondence between language and mental concepts, meaning that the words we use in language directly represent the ideas or meanings in our minds. According to this perspective, when we communicate using language, we aim to represent our thoughts and convey their meaning to others accurately. From this viewpoint, each word or symbol in a language is assigned a specific meaning corresponding to a specific mental concept. When we encounter a word or symbol, we interpret its meaning by associating it with the corresponding mental concept in our minds. This approach assumes that language acts as a transparent medium, allowing for the direct and precise transmission of meaning from one person to another. For example, if we use the word “tree” in a conversation, the representational approach suggests that the word “tree” directly represents the mental concept of a tree

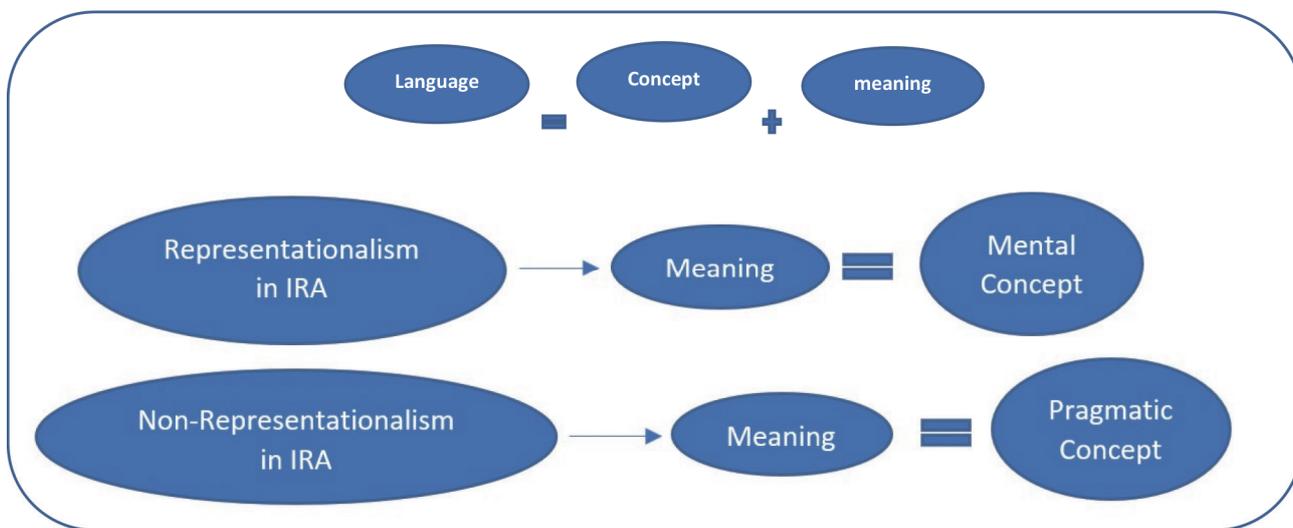


Figure 4. An illustration of Representational and Non-Representational approaches to semantic elements of language (concept + meaning) in IRA

in our minds. Upon hearing the word, the listener would then interpret its meaning by mentally accessing their own concept of a tree and understanding it in the same way as the speaker intended. However, it is important to note that the representational approach does not account for context, culture, or individual differences that can influence the interpretation of meaning. It assumes a one-to-one relationship between words and mental concepts, disregarding the possibility of ambiguity, multiple interpretations, or the influence of subjective experiences and cultural perspectives.

Alternative approaches, such as the pragmatic or social constructionist perspectives, recognize that meaning is not solely determined by a direct representation of mental concepts through language. These approaches consider contextual factors, social interactions, and the shared understanding between individuals as crucial elements in interpreting and negotiating meaning. The representational approach suggests that language directly represents mental concepts, assuming a direct and precise correspondence between words and meaning. However, alternative approaches recognize the influence of contextual and social factors in interpreting meaning, challenging the idea of a one-to-one equivalence between language and mental concepts.

Conversely, the non-representational approach to language suggests that meaning is not solely derived from a direct representation of mental concepts through words or symbols. Instead, it considers meaning as a pragmatic concept. According to this perspective, meaning is not fixed or predetermined but is constructed through a dynamic process involving various contextual and social factors. It recognizes that language is not a transparent medium, and the interpretation of meaning goes beyond a simple mapping between words and mental concepts.

In the non-representational approach, meaning is seen as a result of pragmatic processes such as inference, context, intention, and shared understanding between speakers and listeners. For example, when someone says, "It is hot in here," the non-representational approach considers the interpretation of meaning beyond the literal words. The inference and contextual factors play a role in understanding that the person might be implying a request to adjust the temperature. The non-representational approach also recognizes that cultural norms, social conventions, and individual perspectives can influence meaning. It acknowledges that language use is a social activity, and meaning arises through negotiation between interlocutors based on shared knowledge, context, and communicative intentions. It considers the pragmatic concept of meaning, which focuses on communication's intentions, context, and effects rather than a direct representation of mental concepts. Overall, the non-representational approach to language highlights the significance of pragmatic factors such as context, inference, shared knowledge, and communicative intentions in the interpretation of

meaning, emphasizing that meaning is a dynamic and context-dependent process.

## 5.0 Conclusion

Both IR and KO systems rely on a shared language, necessitating the examination of language-related issues from theoretical and philosophical perspectives. These fields primarily focus on knowledge representation through the semantic elements of language namely meaning and concepts. To comprehensively understand and analyze language representation in KO and IR theories, a deeper exploration of semantic elements of language in the philosophy of language is necessary. This study employed the paradigm of representationalism in philosophy to critically analyze representation of knowledge embedded in the semantic elements of language within a new framework referred to as IRA. It is apparent that representationalism shares certain components with positivism and cognitive science, such as the connection between language and mind, providing an interpretative lens for language users within the context of IRA.

To comprehend language representationally in IRA, it is necessary to delve into semantics, semiotics, and the philosophy of language, which focus on meaning, signs, and concepts. Studies conducted by Blair (1990, 1992, 2003), Mai (1999, 2000, 2001, 2004), Rafferty (2001), Huang (2006), Thornley and Gibb (2007, 2009), and Friedman and Thellefsen (2011) exemplify the significance of these areas in understanding knowledge representation through the semantic elements of language. However, critics of the representational approach to language in IRA argue that it fails to account for language usage's dynamic and context-dependent nature. Consequently, they contend that meaning is not simply a matter of matching predefined symbols to information but rather involves a complex interplay between language, context, and social and cultural factors. As such, they advocate for a more nuanced approach to language and meaning that can fully capture the richness and complexity of human communication. That is why there is a need for further research and discussion on alternative approaches for knowledge representation through the semantic elements of language within the context of IRA, such as non-representationalism, to determine the most effective approach.

In the proposed non-representational approach to the semantic elements of language, meaning and concept are not regarded as mere mental representations. Rather, they encompass pragmatic elements that are subject to modification based on the social practices and behaviors of users within the context of IRA, recognizing that language and meaning are deeply influenced by social context, culture, and community. This approach adopts a pragmatic understanding of language, focusing on how language is used in practice rather than relying solely on formal representa-

tions. It acknowledges that language serves communicative purposes and should be understood in relation to its real-world usage. By considering the pragmatic aspects of language, the non-representational approach aims to enhance the utility and relevance of IR systems within the context of IRA. The dynamic and adaptive nature of language and meaning is another key aspect of the non-representational approach to the semantic elements of language.

Overall, the proposed non-representational approach to the semantic elements of language within the context of IRA offers a holistic and user-centered perspective. While the non-representational approach offers a different perspective from representationalism, it does not necessarily mean abandoning representationalism entirely. Instead, incorporating contextual, pragmatic, and dynamic elements allows for a more comprehensive and nuanced interpretation of meaning and concept. The effectiveness of the non-representational approach for accurate interpretation of meaning and concept within the context of IRA can be evaluated through empirical studies, comparative analyses, and ongoing discussions in the field.

## Endnotes

1. Language, as a semiotic system, encompasses syntax, semantics, and pragmatics. This paper specifically emphasizes the semantic layer, which consists of two elements: meaning and concept.
2. "The overall significance of the sentence -for Frege, its truth or falsehood- is fixed by what the parts of the sentence 'stands for' in the world, and the relations between those parts". (Crane 2005, 575)
3. Frege proposes the 'conceptual representation' versus Russell's 'Referential Theory of Meaning'.

## References

- Blair, David C. 1990. *Language and Representation in Information Retrieval*. Elsevier North-Holland, Inc.
- Blair, David C. 1992. "Information Retrieval and The Philosophy Of Language." *The Computer Journal* 35, no. 3: 200-207. <https://doi.org/10.1093/comjnl/35.3.200>
- Blair, David C. 2003. "Information Retrieval and The Philosophy of Language." *Annual Review of Information Science and Technology* 37, no. 1: 3-50. <https://doi.org/10.1002/aris.1440370102>
- Chu, Heting. 2010. "Information Representation and Retrieval: An Overview". *Information Representation and Retrieval (IRR) in the Digital Age (ASIST Monograph Series)*, 1-21. Medford, NJ: Information Today.
- Crane, Timothy. 2005. "Meaning". In *The Oxford Companion Philosophy*. 2nd ed, edited by Ted Honderish, 575-76. Oxford: Oxford University Press.
- Day, Ronald E. 2016. "Sense in Documentary Reference: Documentation, Literature, And The Post-Documentary Perspective." *Proceedings From the Document Academy* 3, no. 1:6. <https://doi.org/10.35492/docam/3/1/6>
- Ellis, David. 1992. "The Physical and Cognitive Paradigms in Information Retrieval Research." *Journal of Documentation* 48, no. 1: 45-64. <https://doi.org/10.1108/eb026889>
- Egan, Frances 2012. "Representationalism". In *Oxford Handbook of Cognitive Science*, edited by Eric Margolis, Richard Samuels, and Stephen P. Stich, 250-72. Oxford: Oxford University Press.
- Friedman, Alon, and Martin Thellefsen. 2011. "Concept Theory and Semiotics in Knowledge Organization." *Journal of Documentation* 67, no. 4: 644-74. <https://doi.org/10.1108/00220411111145034>
- Frohmann, Bernd. 1990. "Rules of Indexing: A Critique of Mentalism in Information Retrieval Theory." *Journal of Documentation* 46, no. 2: 81-101. <https://doi.org/10.1108/eb026855>
- Hirst, Richard. J. 2006. "Realism". In *Encyclopedia of Philosophy*. 2<sup>nd</sup> ed. Detroit: Thomson Gale Macmillan Reference, 260-69.
- Hjørland, Birger. 1998. "Information Retrieval, Text Composition, And Semantics." *Knowledge Organization* 25, nos. 1/2: 16-31. <https://doi.org/10.5771/0943-7444-1998-1-2-16>
- Hjørland, Birger. 2004. "Arguments for Philosophical Realism in Library and Information Science", *Library Trends* 52, no.3: 488-506.
- Hjørland, Birger. 2007. "Semantics and Knowledge Organization". *Annual Review of Information Science and Technology* 41, no. 1: 367-405. <https://doi.org/10.1002/aris.2007.1440410115>
- Hjørland, Birger. 2011. "The Importance of Theories of Knowledge: Indexing and Information Retrieval as An Example." *Journal of the American Society for Information Science and Technology* 62, no. 1: 72-7. <https://doi.org/10.1002/asi.21451>
- Hjørland, Birger. 2013. "User-Based And Cognitive Approaches to Knowledge Organization: A Theoretical Analysis of The Research Literature." *Knowledge Organization* 40, no. 1: 11-27. <https://10.5771/0943-7444-2013-1-11>
- Hjørland, Birger. 2021. "Information Retrieval and Knowledge Organization: A Perspective from The Philosophy Of Science." *Information* 12, no. 3: 135-60. <https://doi.org/10.3390/info12030135>
- Huang, Sheng-Cheng. 2006. "A Semiotic View of Information: Semiotics as a Foundation of LIS Research in Information Behavior." *Proceedings of the American Society for Information Science and Technology* 43, no. 1: 1-17. <https://doi.org/10.1002/meet.1450430166>

- Ingwersen, Peter. 1992. *Information Retrieval Interaction*. London: Taylor Graham.
- Kelly, Diane. 2019. "What Do We Mean by Theory in Information Retrieval?." In *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval 2-5 October 2019 Santa Clara, United States*, edited by Xi Niu, 83-93. New York: The Association for Computing Machinery.
- Mai, Jens-Erik. 1999. "Deconstructing the Indexing Process." In *Advances in Librarianship 23*, edited by Elisabeth A. Chapman, 269-98. Leeds: Emerald Group Publishing Limited.
- Mai, Jens-Erik. 2000. *The Subject Indexing Process: An Investigation of Problems in Knowledge Representation*. PhD diss., University of Texas at Austin, 2000.
- Mai, Jens-Erik. 2001. "Semiotics and Indexing: An Analysis of The Subject Indexing Process." *Journal of Documentation* 57, no. 5: 591-622. <https://doi.org/10.1108/EUM0000000007095>
- Mai, Jens-Erik. 2004. "Classification in Context: Relativity, Reality, And Representation." *Knowledge Organization* 31, no. 1: 39-48. <https://doi.org/10.5771/0943-7444-2004-1-39>
- Rafferty, Pauline. 2001. "The Representation of Knowledge in Library Classification Schemes." *Knowledge Organization* 28, no. 4: 180-91. <https://doi.org/10.5771/0943-7444-2001-4-180>
- Reese, William. L. 1996. *Dictionary of Philosophy and Religion Eastern and Western Thought*. exp.ed. Pennsylvania State University: Humanity Books.
- Smythe, William. E. 1992. "Positivism and the Prospects for Cognitive Science." In *Positivism in Psychology: Historical and Contemporary Problems*, edited by Charles W. Tolman, 103-18. New York: Springer-Verlag.
- Stock, Wolfgang G. 2010. "Concepts and Semantic Relations In Information Science." *Journal of the American Society for Information Science and Technology* 61, no. 10: 1951-69. <https://doi.org/10.1002/asi.21382>
- Svenonius, Elaine. 2004. "The Epistemological Foundations of Knowledge Representations". *Library Trends* 52, no. 3: 571-87.
- Tye, Michael. 2002. "Representationalism and the Transparency of Experience." *Noûs* 36, no. 1: 137-51. <https://doi.org/10.1111/1468-0068.00365>
- Thornley, Clare and Forbes Gibb. 2007. "A Dialectical Approach to Information Retrieval." *Journal of Documentation* 63, no. 5: 755-64. <https://doi.org/10.1108/00220410710827781>
- Thornley, Clare and Forbes Gibb. 2009. "Meaning in Philosophy and Meaning in Information Retrieval (IR)." *Journal of Documentation* 65, no. 1:133-50. <https://doi.org/10.1108/00220410910926158>
- Weiss, Leila C., Marisa Bräscher, and William B. Vianna. 2016. "Pragmatism, Constructivism and Knowledge Organization." In *Knowledge Organization for a Sustainable World, Proceedings of the Fourteenth International ISKO Conference, 27-29 September 2016 Rio de Janeiro, Brazil* edited by José Augusto Chaves Guimarães, Suelen Oliveira Milani and Vera Dodebei. Advances in Knowledge Organization 15. Würzburg: Ergon Verlag, 211-18.
- Yang, Dachun. 2008. "Representationalism and The Linguistic Question in Early Modern Philosophy." *Frontiers of Philosophy in China* 3, no. 4: 595-606.