

Martin Thellefsen, Torkild Thellefsen, bent Sorensen

Introducing a semeiotic-inspired concept of information for library and information science – breaking the boundaries of documents

Studia Kulturoznawcze nr 1 (7), 267-280

2015

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

MARTIN THELLEFSEN, TORKILD THELLEFSEN

Royal School of Information and Library Science
Copenhagen University

BENT SØRENSEN

Independent Scholar

Introducing a semeiotic-inspired concept of information for library and information science – breaking the boundaries of documents

The consciousness of a general idea has a certain “unity of the ego” in it, which is identical when it passes from one mind to another. It is, therefore, quite analogous to a person, and indeed, a person is only a particular kind of general idea.

Charles Sanders Peirce,
Man’s Glassy Essence, *The Monist*, Vol. III, 1/1892

Introduction

Library and information science (LIS), and in particular its subfield Knowledge Organization (KO), has always had a particular interest in the development of better systems for information retrieval through documents, while at the same time maintaining a particular focus on users’ information behavior. The very *raison d’être* and uniqueness of LIS lie in its foundation on the complex intersection between information systems and the users, consumers or recipients using those systems. The developments of LIS may roughly be described in terms of different paradigmatic stages, from the early online age in the fifties, also identified as the systems-oriented paradigm, through a user-oriented cognitive paradigm, and in particular the cognitive view of the nineties, up to the present time, often described as the postmodern age, which offers a vast range of interactive technologies and possibilities for users and consumers and their insatiable thirst for information. Looking back through the decades of LIS,

we may note that the development of computers and information technology has always been a driving force that seems to be constantly accelerating and transforming the landscape of the information society. However, it is important to realize that even though technology and computer science are important driving forces, it may only be so because of the innovative and creative uses they motivate. Consequently, what really motivates changes in the information environment is not technology itself, but how it is used to accommodate the information tasks of users. Even though these three stages are often conceptualized as different paradigms, they remain as co-existing views or schools within LIS. In principle, even though there is little knowledge exchange between the three views, as demonstrated for example by bibliometric mappings (White and McCain, 1998; Åström, 2007), they are deeply interdependent.

The concept of information is undoubtedly essential within LIS; however, it seems that the three above-mentioned paradigms operate with different conceptions of information. What really connects the three paradigms is the document. It is documents that make up the pool of information in LIS, out of which all types of research spring, from user behavior studies to bibliometric studies, to knowledge organization, etc. Being such an important concept, the document has of course been the subject of vast amounts of study within LIS, from the perspectives of knowledge organization, knowledge communication and information retrieval (see for example: Buckland, 1997; Fjordback, Andersen and Hjørland, 2003; Frohmann, 2004, 2008; Ingwersen, 2002; Järvelin and Kekäläinen, 2000; Mai, 2005; Olsen, Lund, Ellingsen and Hartvigsen, 2012).

The aim of this paper is to demonstrate: 1) that the three different views articulate different but not mutually exclusive concepts of information; 2) that these different concepts of information, co-existing within the rather small research field that LIS represents, are tied to different kinds of information praxis or behavior, such as searching and retrieval; 3) that these three concepts of information are linked to documents; and 4) that semeiotic can provide a useful framework that enables us to see information within these three views as different levels of signification, and furthermore, provide us with a concept of information that transcends the boundaries of the document and thus enables us to address information more generally within LIS without reference to documents and documentation.

1. Describing the three dominant concepts of information in LIS

The concept of information has been discussed in detail by scholars within LIS, and in particular Capurro and Hjørland (2003) give a thorough review of the

concept of information, its historical developments and its different definitions and applications in LIS. We do not wish to reproduce this discussion, but merely refer the interested reader to this excellent review article. Our objective with this section is, using the triadic semeiotic model, to summarize the three fundamental views on information, and demonstrate how they are different in perspective, and as such, relate to different epistemological presuppositions.

2. The systems-oriented view and its concept of information

The systems-oriented perspective emerged with the online era in the 1950s, and is best described with reference to the canonical Cranefield studies in information retrieval conducted by Cyril W. Cleverdon (1960). The developments in automatic indexing, also used by today's Internet search engines, can be traced back to the pioneering work done by Cleverdon. The systems-oriented view operates with an objective concept of information and relevance that is close to a datalogical framework, and consequently it is often described as the information-processing paradigm (Brier, 2004). This leaves us with a concept of information that is based not on ontological presuppositions but rather on a positivist epistemology – information is measurable at different levels of granularity. In this perspective, the concept of information relates only to computers' information processes. If we apply a triadic analysis to the systems-oriented perspective, it may be presented in the following way:

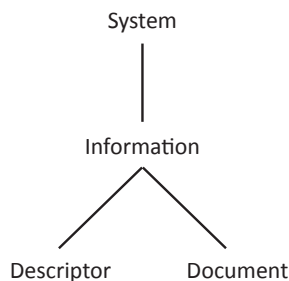


Figure 1. The information concept of the systems-oriented perspective is made up of relations between documents on the one hand (the data pool, so to speak) and descriptors on the other, mediated by the system

3. The user-oriented view and its concept of information

The user-oriented view fundamentally acknowledges the research carried out within the framework of computer science (and the systems-oriented view),

but has a particular focus on the user's information-seeking behavior, information needs and interaction with the information system. Within this view, the user's cognitive state and thus motives for seeking information in a database (whether bibliographical or encompassing the Internet in general) is of particular interest. In this perspective too, information is objective; however, the relevance judgment is increasingly considered subjective and linked to a particular task (Borlund, 2009).

In the user-oriented view, as defined by Nicholas Belkin (1978), information must be purposeful, meaningful and new; it must enter into human communication. It must be requested and desired. It must have an effect on the recipient. It must affect the state of knowledge of the recipient. It must be general and able to predict the effect of information. Consequently, information is only information insofar as a generator is capable of communicating the information in a purposeful and meaningful way. But this is only one side of the equation; on the other side, the recipient's side, the recipient must request or desire information, and given this desire or request, the information must have an effect on his state of knowledge. Information, even though related to the information processing paradigm (as described above), is also related to subjectivist epistemology (Hjørland, 2002). In this perspective, the interest in information relates solely to the interaction between a user's information-seeking and the particular information system. Following the triadic model of semeiotic, this can be presented in the following model.

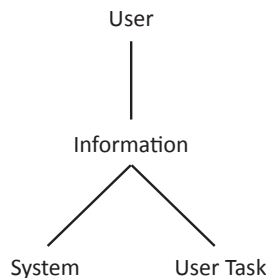


Figure 2. The information concept of the user-oriented perspective is a relation between user tasks on the one hand and the system on the other, mediated by the user's relevance judgment

4. The domain-oriented view and its concept of information

The domain-oriented view (or discourse-oriented view) takes a different stance than the two above-mentioned views. Information in this perspective cannot

be reduced to objective representations and data structures, or to subjective or personal information needs, but must be defined pragmatically and paradigmatically, and thus involves aspects of situation, context and knowledge interests framed within a discourse community. In this case, neither information nor relevance judgment is objective. Birger Hjørland, the founder of the domain-oriented view (DA: Domain Analysis) argues however that information should be replaced with documentation:

“Information science” developed out of documentation, and the documentation movement developed a theoretically motivated concept “document” as a basic term [...] for the field: A document is “any concrete or symbolic indication, preserved or recorded, for reconstructing or for proving a phenomenon, whether physical or mental” (Brier, 1951: 7; quoted from Buckland, 1991).

In this perspective information becomes equal to documentation. This view can be presented in the triadic semeiotic model in the following way:

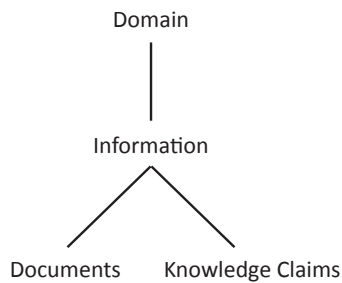


Figure 3. The information concept of the domain-oriented perspective is made up of the relation between knowledge claims on the one hand and documents on the other, mediated by the domain

Based on this short analysis, we believe that LIS would profit from a concept of information that is theoretically robust and consistent at all levels of signification, and furthermore at a meta-theoretical level can provide for a medium-independent conception of information (by medium we mean, for example, the document). Consequently, we are not trying to establish a framework for merging the three views, but to propose a theoretical foundation of information based in semeiotic. Simply put, we consider information as signs that operate at different levels of granularity depending on the sign generator, the sign user and the context of the signification process.

5. The semeiotic-oriented view and its concept of information

As stated above, we propose a semeiotically inspired concept of information; however, paraphrasing Peirce's distinction between immediate and dynamical object (EP 2: 498), we define information as either passive (primarily iconic and indexical – as degenerate signs) or active (as genuine symbols). Passive information is information in documents. This information simply resides in documents passively. Passive information is information which is not in interpretative use. This is also what we would describe as the ontological level of information; it has a being in itself. Active information, on the other hand, is information being interpreted; this is information taking part in a signification or communication process. This can be called the epistemological level of information. As stated above, we consider a semeiotic-inspired concept of information independently of a particular medium, thus including any act of communication. Consequently, our focus is not documents or documentation, but information at all levels of signification. Therefore, a semeiotic concept of information is able to address information as signs independent of documents; information that exists between actors in a knowledge domain as a form of collateral experience, which according to Peirce is:

All that part of the understanding of the Sign which the Interpreting Mind has needed collateral observation for is outside the Interpretant. I do not mean by "collateral observation" acquaintance with the system of signs. What is so gathered is *not* COLLATERAL. It is on the contrary the prerequisite for getting any idea signified by the sign. But by collateral observation, I mean previous acquaintance with what the sign denotes (EP 2: 494).

Furthermore, our concept of information addresses information as knowledge that is anchored in the epistemology and the values of the knowledge domain. These types of information are not explicitly documented, and therefore cannot be identified using ordinary methods within KO (this has been widely discussed in Thellefsen, 2002, 2004, 2010; Thellefsen and Thellefsen, 2004). However, before we address this important theme, we will try to lay bare our concept of information.

We understand all information to be signs of the type: icon, index and symbol, following the semeiotic terminology of Charles S. Peirce. This means that information either resembles its object, or points to its object or represents its object through convention.

According to Peirce, an Icon is a sign "which represents its object by virtue of a character which it would equally possess did the object and the interpreting

mind not exist" (MS [R] 142: 3–4). The icon stands in a passive relation to its object; it just happens that it shares some qualities with an object. However it is an important part of the symbol, since it is what makes us capable of distinguishing between objects: "the idea embedded by an icon [...] cannot of itself convey any information, being applicable to everything or nothing" (CP 3: 433). We may add, however, that hypo-icons are able to convey information (Nöth, 2012).

The index is a sign "which represents its object by virtue of a character which it could not possess did the object not exist, but which it would equally possess did the interpreting mind not operate" (MS [R] 142: 3–4). The index is what points out the symbol in time and place and connects our thoughts to reality (Nöth, 2014). Peirce further writes that "The index asserts nothing; it only says "There!" It takes hold of our eyes, as it was, and forcibly directs them to a particular object, and there it stops" (CP 3: 361).

The symbol is a sign which represents its object by virtue of a character which is conferred upon it by an operation of the mind (MS [R] 142: 3–4).

Nöth sums up the icon and index components in symbols in the following way: "Icons are needed to show what we are talking about and indices to connect our thoughts to the reality which they represent. Symbols are associated to the objects they represent by habits" (Nöth, 2014: 177).

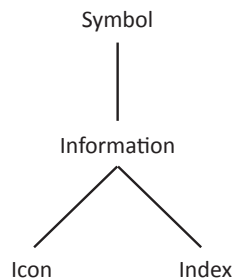


Figure 4. The information concept of the semeiotic perspective is made up of the index on the one hand and the icon on the other, mediated by the symbol

However, when information is one of these signs (symbols always involve indices and icons; indices always involve icons) it exists as passive information at an ontological level because it does not involve interpretation in order to exist; it may have been active information, before it became passive at the moment it no longer took part in a signification or communication process, and it may be active again. It is only when information is communicated or interpreted that it becomes active; it then becomes signs of the type rheme, dicisign or argument, still following Peirce's classification. It is worth noting that the last-mentioned signs are also named rhema, proposition and argument (EP 2:

204). An argument must involve propositions and rhemae, and propositions must involve rhemae. Consequently, we can say that information at an epistemological level must take part in communication as propositions and it always involves symbols (which involves indices and icons). Following this logic, there cannot be any epistemological level of information without an ontological ditto. Passive and active information are interdependent. Below we have attempted to schematize the two levels of information.

Information

Ontology (passive)	Epistemology (active)
Icon	Rhema
Index	Proposition
Symbol	Argument

The point is that information ontologically can exist outside the mind, and its existence does not depend on pragmatic factors such as a sender or a receiver, nor does it involve intentionality, purpose or being desired. At the epistemological level information depends on these pragmatic factors. To further describe this two-sidedness (passive, active) of information, we apply the concepts of ego vs. non-ego. Peirce describes the idea in the following way:

It is an experience. It comes out most fully in the shock of reaction between ego and non-ego. It is there the double consciousness of effort and resistance. That is something which cannot properly be conceived. For to conceive it is to generalize it; and to generalize it is to miss altogether the *hereness* and *nowness* which is its essence (CP 8: 266).

Non-ego is information, whether it be passive or active, it is outside the mind; ego is the sum of our experiences and it is our ability to cognize what mediates between the outer and inner world. It is information that forms the ego (which is always in process) following the dictum of Aristotle, nothing is in the intellect which was not first in the senses. The ego is made up from experience caused by information, organized as knowledge through cognition. This can be schematized in the following way.

Information	
Ontology (passive)	Epistemology (active)
Ego	non-ego
Qualisign	icon
Sinsign	index
Legisign	symbol
	rhema
	proposition
	argument

By applying this two-sidedness of information we suggest an interdependent relationship between the ontological and epistemological levels of infor-

mation. In fact, we suggest that this interdependent relation is a necessary condition in any meaning creation process where information enters the mind through cognition. Let us elaborate on this perspective in Figure 5.

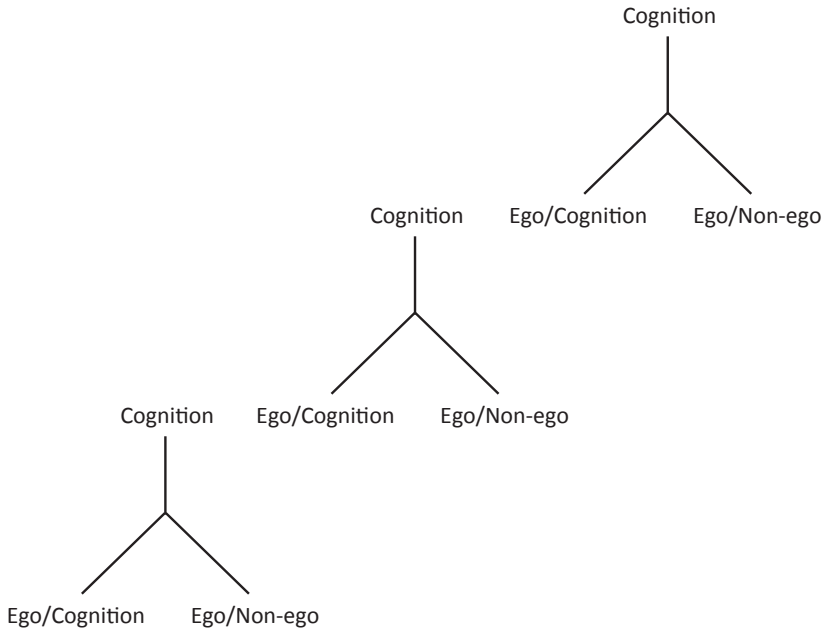


Figure 5. The dynamics of meaning creation

We suggest that information, understood as non-ego, exists outside ego (the mind) and information enters the mind via perception; it then becomes interpreted, related to our background knowledge and transformed into knowledge – this is an act of cognition. When the information, now as knowledge, is communicated, it becomes information again. The logic of the diagram is that non-ego becomes ego mediated by cognition. It is through cognition that information enters the mind – and we use cognition in a very broad sense, also including what lies beneath the self-consciousness. Consequently, ego (now with the interpreted information) becomes ego in the next triad and so on and so forth. Cognition can only happen when the ego receives information – Figure 5 suggests this dynamic process.

Figure 6 (below) elaborates on Figure 5. Triad A is information on an ontological level. Triad B is our consciousness in semeiotic terms. According to this triad, consciousness involves three elements: primisense, altersense and medisense; without going too deeply into this terminology, they make up our consciousness, which involves an inner world (primisense, which is ego), an

outer world (altersense, which is non-ego) and mediation (medisense, which is cognition). In an untitled manuscript (c. 1900), Peirce wrote:

There are no other forms of consciousness except [...] Feeling, Altersense, and Medisense. They form a sort of system. Feeling is the momentarily present contents of consciousness taken in its pristine simplicity, apart from anything else. It is consciousness in its first state, and might be called primisense. Altersense is the consciousness of a directly present other or second, withstanding us. Medisense is the consciousness of a thirdness, or medium between primisense and altersense leading from the former to the latter. It is the consciousness of a process of bringing to mind (CP 7: 551).

It is when medisense mediates between primisense and altersense that information becomes knowledge; thus we are at an epistemological level. Figure 6 is a classification of associations following the semeiotic logic of three. Every interpretation of information will lead to one or more associations of either similarity (iconicity), contiguity (indexicality) or interest (symbolicity). This also takes place at an epistemological level.

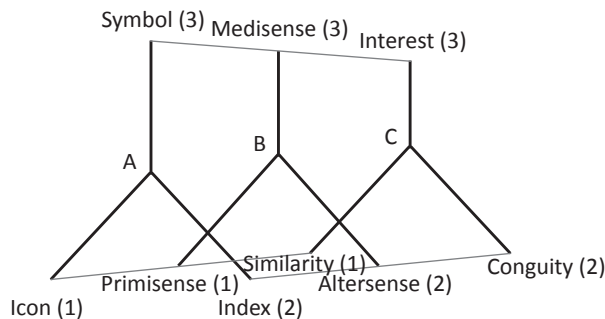


Figure 6. Integration of signs, types of consciousness and associations

We believe that the concept of information presented here encompasses information on an ontological as well as an epistemological level. Knowledge about the ontological level of information is necessary since it determines how we ought to work with information. Epistemology cannot grasp information alone, since information will slip through our fingers and disappear before we can describe it. We need to know what our knowledge is about.

How can this semeiotic-inspired concept of information also encompass the three concepts of information in KO which were discussed earlier?

As shown in the triads under each perspective, we believe that there is a semeiosis or a process of signification taking place within the computer. There is an indexical relation between the documents in the database and the descrip-

tors describing the documents. The system is the mediator – it is a quasi-interpreter or simply and solely a symbol. The user-oriented perspective involves a semeiotic relation between a given task performed in a system, and it is the relevance of the retrieved documents as icons that the interpreter interprets as information. In the domain-oriented perspective a third element is added – a domain, which is the context in which the knowledge claims are identified in documents. It seems that our semeiotic information concept is capable of handling the three perspectives, as it can involve the processes within the computer, the relation between a user and a system, and the knowledge within a knowledge domain. Thus, besides being able to encompass the three perspectives, we also believe that our concept of information can break the stronghold of documents. This means that the information we use in order to perform (for example) knowledge organization does not take its starting point in documents alone, but rather in the epistemological basis of the knowledge domain or the values of a community. The epistemological basis of a knowledge domain is the sum of choices made by the knowledge domain in relation to the object(s) being investigated (Thellefsen, 2002, 2004, 2010).

6. Strong and weak ontologies

One may ask why it is necessary to introduce a fourth perspective when there are three already. What can the semeiotic perspective contribute that is not already in the three existing perspectives? In order to point toward answers to these questions, we have constructed the following table, showing how signification and communication in the four perspectives involve different levels of ontology, passing from weak to strong ontologies. When it comes to the three perspectives, they all have weak ontologies, which means that they – at best – have ideas about reality as something that in some undefined or unreflected way puts indirect pressure on signification and communication and also information; however, these ideas are not systematically elaborated. This leads to the Table.

Apart from the semeiotic perspective, all of the perspectives have weak ontologies, which means that they do not have outlined theories of signification or communication that include information having an existence in itself. The semeiotic perspective, on the other hand, has a strong ontology, which means that information exists outside signification and communication.

The reason for us to address an ontological level of both signification and communication is to break the stronghold of documents, which dominates the three perspectives. We propose a concept of information which is involved in both signification and communication, and is not tied to documents – rather

Table 1. Strong and weak ontologies in signification and communication

Perspective	Signification		Communication	
	Strong ontology	Weak ontology	Strong ontology	Weak ontology
Systems-oriented perspective		x		x
User-oriented perspective		x		x
Domain-oriented perspective		x		x
Semeiotic perspective	x		x	

our semeiotic information concept makes information possible everywhere; the entire universe is perfused with information. In this way we can handle information outside documents, this will make our concept of information more suitable for the future than the three other perspectives. Also our semeiotic concept of information is able to encompass the three dominant perspectives in a signification context. In the relation between the system, the user and the domain, there are numerous processes of signification taking place. However, being fragmented, these perspectives do not take this overview into consideration. Our semeiotic perspective is able to understand the content of the computer as a potential, the particular search of the user as an actualization of this potential, and the understanding of the result of the search in terms of interpretation as the domain; in this manner we do not combine the three perspectives, but we pinpoint the relation between the system, the user and the domain; and because of our strong ontology we are able to use the concept of sign within all perspectives.

References

- Annotated Catalogue of the Papers of Charles S. Peirce (MS) (1967)*, ed. by R.S. Robins, Boston: University of Massachusetts Press (microfilm edition of the manuscripts of C.S. Peirce).
- Borlund, P. (2009). User-centered evaluation of information retrieval systems. In: A. Göker, J. Davis (eds.). *Information Retrieval: Searching in the 21st Century* (pp. 21-36). London: John Wiley & Sons.
- Brier, S. (2004). Cybersemiotics and the problems of the information-processing paradigm as a candidate for a unified science of information behind library information science. *Library Trends*, 52(3), pp. 629–657.
- Belkin, N.J. (1978). Information Concepts for Information-Science. *Journal of Documentation*, 34(1), pp. 55–85.
- Collected Papers of Charles S. Peirce (CP) (1931–1935)*, Vol. 1–8, ed. by Ch. Hartshorne, P. Weiss, Cambridge, Mass.: Harvard University Press.

- Capurro, R., Hjørland B. (2003). The concept of information. *Annual Review of Information Science & Technology*, 37, pp. 343–411.
- Fjordback, T., Andersen, J., Hjørland, B. (2003). Documents and the communication of scientific and scholarly information: Revising and updating the UNISIST model. *Journal of Documentation*, 59(3), pp. 278–320.
- Frohmann, B. (2004). Documentation redux: Prolegomenon to (another) philosophy of information. *Library Trends*, 52(3), pp. 387–407.
- Frohmann, B. (2008). Revisiting “What is a document?”. *Journal of Documentation*, 65(2), pp. 291–303.
- Hjørland, B. (2002). Epistemology and the Socio-Cognitive Perspective in Information Science. *Journal of the American Society for Information Science and Technology*, 53(4), pp. 257–270.
- Ingwersen, P. (2002). Cognitive perspectives of document representation. *Colis4: Emerging Frameworks and Methods*, pp. 285–300.
- Järvelin, K., Kekäläinen, J. (2000). *IR evaluation methods for retrieving highly relevant documents*. Proceedings of the 23th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (ACM SIGIR '00), 2000, July 24–28. Athens.
- Mai, J.E. (2005). Analysis in Indexing: Document and Domain Centered Approaches. *Information Processing and Management*, 41(3), pp. 599–611.
- Nöth, W. (2012). Medieval maps: Hybrid ideographic and geographic sign systems. In: I. Baumgärtner, M. Stercken (eds.). *Herrschaft verorten: Politische Kartographie des Mittelalters und der Frühen Neuzeit* (pp. 335–353). Zürich: Chronos.
- Nöth, W. (2014). The growth of signs. *Sign Systems Studies*, 42(2/3), pp. 172–192.
- Olsen, B., Lund, I., N.W., Ellingsen, G., Hartvigsen, G. (2012). Document theory for the design of socio-technical systems: A document model as ontology of human expression. *Journal of Documentation*, 68(1), pp. 100–126.
- The Essential Peirce. Selected Philosophical Writings* (EP), Vol. 1: 1867–1893, ed. by N. Houser, Ch. Kloesel (1992), Vol. 2: 1893–1913, Peirce Edition Project (1998). Bloomington: Indiana University Press.
- Thellefsen, T. (2002). Semiotic Knowledge Organization: Theory and method development. *Semiotica*, 142(1/4), pp. 71–90.
- Thellefsen, T. (2004). Knowledge Profiling: The Basis for Knowledge Organization. *Library Trends*, 144, pp. 507–514.
- Thellefsen, T. (2010). *A Semiotic Outline of Significance-effect, Fundamental Sign and Knowledge Profiling and Their Use in Knowledge Organization and Branding*. Saarbrücken: VDM Verlag Dr. Müller.
- Thellefsen, T., Sørensen, B., Thellefsen, M. (2014). The Information Concept of Nicholas Belkin Revisited – some Semeiotic Comments. *Journal of Documentation*, 70, pp. 74–92.
- Thellefsen, M., Sørensen, B., Thellefsen, T. (2015). The concept of information in Library and Information science – a field in search of its boundaries – 9 short comments concerning information. *Cybernetics and Human Knowing* (in press).
- Thellefsen, T., Thellefsen M. (2004). Pragmatic Semiotics and Knowledge Organization. *Knowledge Organization*, 31, pp. 177–187.

- Thellefsen, T., Thellefsen, M., Sørensen B. (2013). Emotion, information and cognition and some possible consequences for Library and Information Science. *Journal of the American Society for Information Science and Technology*, 64, pp. 1735–1750.
- White, H. D., McCain, K. W. (1998). Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science*, 49(4), pp. 327–355.
- Åström, F. (2007). Changes in the LIS research front: Time-sliced cocitation analysis of LIS journal articles, 1990–2004. *Journal of the American Society for Information Science and Technology*, 58(7), pp. 947–957.

Summary

The field of Library and Information Science (LIS) suffers from a fragmented view of information. Within LIS there are three dominant perspectives, operating with three different concepts of information. This paper introduces a semeiotic-inspired concept of information, which is able to replace the other concepts of information in LIS, offering a unifying concept which can explain the semeiotic processes that take place when dealing with information. The semeiotic-inspired concept of information is anchored in a strong ontology, involving Peirce's three important signs: icon, index and symbol.

Słowa kluczowe: sem(e)iotyka, informacja, bibliotekoznawstwo i informacja naukowa, znaki, Peirce

Keywords: semeiotics, information, library and information science, signs, Peirce