Article

em questão
E-ISSN 1808-5245

Some core concepts of domain analysis

Birger Hjørlan
University of Copenhagen, Copenhagen, Denmark;
birger.hjorland@hum.ku.dk; https://orcid.org/0000-0001-9535-1732

Abstract: This article claims that information systems should provide reliable and trustworthy information, and that this demand imply a re-orientation of information science with knowledge organization from, for example, approaches based on studies of user's preferences, towards studies of quality and the values and goals in the documents being mediated. This mediation is done by information professionals in their knowledge organization systems and –processes. For this purposes, the philosophy of science becomes an important subject, which has implications for the understanding of a long range of concepts. Here, the following concepts are briefly introduced: "social epistemology", "information", "question," "relevance," "information need," "concept," and "subject" (of documents).

Keywords: domain analysis; social epistemology; information need; information

1 Introduction

A core principle in domain analysis is that information systems should provide reliable and trustworthy information, corresponding to what Furner (2018) called "a veritistic turn." By implication, information science with knowledge organization (KO) is not primarily about doing marketing research, in order to adapt the knowledge organizing systems (KOS) to the users' preferences. Rather, KO is but about securing the authoritativeness of the documents, information and knowledge provided to the users. (User studies may serve a secondary role, for example, by showing that many users prefer verbal systems to systems using classification codes). This view has often been recognized in KO, although there are important exceptions, such as the *Dewey Decimal Classification* (DDC).[1] A prominent library classificationists in favor of providing trustworthy classifications was Henry Bliss, who attempted to base

classification on the scientific and educational consensus.[2] Other parts of library and information science (LIS) has not shared this view. This is, in particular, the case with research in "information behavior," which to a higher degree has been influenced by cognitive and user-oriented views focusing on user preferences rather than on the trustworthiness of information. Research in information behavior is therefore closer related to market research than is KO.

Such a "veritistic turn" in information science and KO have implications of our overall approach to research and practice. Moreover, it has implications for a long range of basic concepts in the fields. Below the following concept are briefly introduced from the domain analytic point of view: Section 2: social epistemology (SE), Section 3: information, Section 4: question, Section 5: relevance, Section 6: information need, Section 7: concept, and Section 8: subject" (of documents). Section 9 provides a general conclusion.

2 The concept "social epistemology"

We should be very proud, that social epistemology (SE) was first introduced by Jesse Shera (1951, p. 82), a researcher in LIS. He wrote (italics added):

> Even a cursory examination of the history of classification of the sciences emphasized the extent to which any attempt to organize knowledge is conditioned by the *social epistemology* of the age in which it was produced. […]. Here, then, is an implicit denial of Bliss' faith in the existence of a 'fundamental order of nature', a rejection of the belief that there is a single, universal, logically divided classification of knowledge.

SE means that the individual's knowledge and thinking is shaped by concepts developed in the tradition in which he or she was socialized and trained. Classifications, for example, are made by people, influenced by certain views and theories. It means, with the words of Fleck ([1935] 1979, p. 38):

> [C]ognition must not be construed as only a dual relationship between the knowing subject and the object to be known. The existing fund of knowledge must be a third partner in this relation as a basic factor of all new knowledge. […]

> Cognition is therefore not an individual process of any theoretical 'particular consciousness'. Rather it is the result of a social activity, since the existing stock of knowledge exceeds the range available to any one individual.

SE is not an alternative to empirical studies or to logical thinking, but depends on them. It is, however, an alternative to empiricisms and rationalisms, which, as individualist epistemologies, neglects the historical, social and cultural influences on the individual's perception and cognition. An empiricist may say: "It is a fact, that I saw a blackbird." However, according to SE, your concept "blackbird" is something that you have learned by somebody informed by some theory on how to classify birds. According to new research based on DNA-analysis, even "blackbird" turns out to be an ambiguous concept (Fjeldså 2013, p. 141).[3] The most important consequences of SE is that it makes subject theories important for all our knowledge claims. Domain analysis is a further development of Shera's SE (see further in Hjørland 2024).

3 The concept "information"

The term "information" is a polysemic term, with different meanings in Shannon's information theory, in the cognitive approach in LIS, in the semiotic view and domain-analytic view etc. If we consider "information" to be anything that may inform somebody about something, then, with Buckland (1991, p. 50), we must conclude, "[…] *we are unable to say confidently of anything that it could not be information*". The implication is that information systems and services need to consider not just the contents of messages, but also the problem and question, which the information support. In principle, the same document may contain different information and needs to be indexed or classified differently in different contexts.

The term "information retrieval" (IR) should be considered a misnomer for "document retrieval". As Lancaster (1968) wrote: "An information retrieval system does not inform (i.e., change the knowledge of) the user on the subject of his inquiry. It merely informs on the existence (or non-existence) and whereabouts of documents relating to his request".

Only one serious attempt going against Lancaster's view and defending the term IR has been identified; this is van Rijsbergen and Lalmas (1996), who wrote: "We believe that the purpose of an information retrieval system is to provide information about a request and that a request is a representation of an information need that an IR system attempts to satisfy."

Rijsbergen and Lalmas' (1996) article is not convincing and is based on a problematic individualistic view. A layperson may consider, for example, some medical information relevant, while experts consider it non-relevant, and vice versa: the layperson may fail to see some information as relevant, which the expert considers relevant (experts may disagree too, of course). Therefore, it is not the isolated individual's judgement, but the collectively established knowledge, which must be the point of departure. This demands that documents, rather than information, be retrieved, because the authority, methodology, etc. of the documents must be taken into account. A match between terms in a query and a document is not sufficient to consider it a fulfillment of the user's information need (not even if the terms have been translated by, for example, a thesaurus). (See also Hjørland, 2023, p. 65, note 37). There are systems, such as ChatGPT, which provides answers/information rather than documents. However, such answers cannot be trusted, because we need to know on what basis they are provided, and even when ChatGPT provides sources, they are often deficient. The real challenge is to evaluate the authority of different documents when using them for answering questions.

4 The concept "question"

Kwon (2016, p. 2) argued that:

> If 'information' is a central concept for library and information science, then 'questions' are fundamental, for information 'informs' relative to the question. But research focusing on questions as a central theoretical concept has been stymied by the paradox of the question, which observes that in order to ask one must know enough to know what one does not know (Flammer, 1981). This dissertation proposes that this paradox results from the limitations of the cognitive approach to questions as indications of individual information need, and that the paradox can be resolved by reframing

> questions as social epistemological tools of inquiry within knowledge domains.

Kwon (2016) shows that research on questioning in information science has conceptualized a question primarily in two ways:

a) as an object that can be conveyed through information transfer; and as a sign of an information seeker's "problematic situation," in which the question is taken at face value, hereby overlooking the complexity of questioners' cognitive processes of formulating and asking questions;
b) focusing on questioning as an individual's mental activity, on tactics for clarifying people's queries to get at the "correct" information to respond to their "visceral need" (Taylor, 1968). Here questions are considered residing and developing purely within individuals, neglecting the SE perspective of questions formed on the basis on some background assumptions about theories, concepts and knowledge in the domain of the question;

Kwon considers both strategies as problematic and suggests conceptualizing questioning in a third way;

c) as an intersubjective practice, wherein questions are construed as practices rooted in social norms within knowledge domains (see also Gorichanaz, 2019).

When a student suggests a research question to a university professor, or when she poses a search suggestion to an experienced information specialist, the question is not taken for granted, but is negotiated with the student. In the cognitive tradition, this negotiation is influenced by Taylor's (1968) view to find the real question inside the student's mind. In domain analysis, the negotiation instead takes it point of departure about the subject literature in the field (e.g., is the question too broad in relation to what has been published? Is it based on scholarly conceptions, terminology and knowledge? Does it entail a perspective that makes it possible to manage the large literature in the field? etc.).

5 Relevance

When we search for documents, we would like to find those that are relevant for us, and avoid those that are not relevant (but represent "noise"). This led to the well-known measures of "recall" and "precision," where KOS are evaluated by their ability to retrieve all relevant documents, while minimizing the number of non-relevant documents. The most prolific author about relevance in information science is Tefko Saracevic, who, in 1975 (p. 333; italics in the original), wrote a much-cited article in which he declared "the subject knowledge view" to be the most fundamental perspective of relevance:

> I wish to suggest that *the subject knowledge view of relevance is fundamental to all other views of relevance*, because subject knowledge is fundamental to communication of knowledge. In that lies the importance and urgency of the work on that view.

This 1975 quote expressed an important view. However, as discussed in Hjørland (2010, p. 217), Saracevic, despite its declared urgency, never since took this view up again. Rather he limited his scope to user-oriented views, and seems not since to have improved our understanding of relevance. Rather, he often declared relevance a badly understood concept or even a "headache," [4] a view with which White (2017, p. 3937) disagreed "Writers on relevance in IS often observe that the concept is not well understood and call for further research to clarify and ultimately measure it.[5] The conclusions here are that, on the contrary, we understand relevance quite well".

I find that White is right, and that the headache is due to a wrong "paradigm" by many researchers in information science, because they have sought for relevance in the wrong place: in the mind of individual users. Soergel (1976, p. 257) and Fugmann (1973) are among the early researchers, who realized the problems in user-based relevance assessment.

Relevance researchers have so far never spoken about how they themselves determine which documents they consider relevant for their research. If they did so, they would discover that relevance assessments depends on theory and "paradigm". In knowledge organization, for example, people subscribing to the facet analytic methodology tend to cite Ranganathan (because

they find his writing relevant) while people subscribing to the domain analytic view tend to cite Hjørland. The same is the case in any field of knowledge, and this is rather easy to understand, but nonetheless not recognized by most of the authors writing about this concept in LIS.

The concept "relevance" is thus easy to understand: A medical treatment is relevant if it has a high probability to cure patients with a given disease and have no serious side effect. The same understanding can be used about documents and information: A document is relevant for a query about the cure for a disease if it provide evidence about the effect of the cure on that disease. Therefore, Hjørland and Sejer Christensen (2002, p. 964) suggested the following definition of the term "relevance" "Something (A) is relevant to a task (T) if it increases the likelihood of accomplishing the goal (G), which is implied by T".

That the concept is easy to understand must not be confused with the problem to determine whether a specific thing or document is relevant to a specific problem.[6] Everybody can grasp the definition, but the specific evaluation of what is relevant requires adequate subject knowledge: You can understand what it means for a medical cure to be relevant, but to determine whether a specific cure is relevant in a specific case, mostly requires medical knowledge. Relevance criteria are therefore developed and discussed on a collective level, and not determined by the individual user.

It is remarkable, as Saracevic (2012, p. 51) pointed out, that relevance has only been considered in relation to searching, never in relation to classification and indexing.[7] According to the domain analytic view, classification and indexing is not just about being right, but also about considering relevance.

6 User need

The concept "information need" was defined by Case and Given (2016, p. 371, glossary):

> Information need: A hypothesized state brought about when an individual realizes that they are not comfortable with their current state of knowledge. According to Taylor (1968), an information need may go through the following four stages on the way to being articulated: visceral, unconscious, formalized, and compromised.

Case and Given's (2016) as well as Taylor's (1968) theory about information needs are typically based on an individualistic rather than on a social epistemology. Taylors most obviously so, as if such needs exists deep inside the individual, and gradually grows up until they, in a compromised way, can be communicated to, for example, a librarian, who then may help satisfying them. This is done, found Taylor, by tracing the need back to the individual's real need by peeling away the social layer in the formulation of the question.
In Hjørland (2010, p. 222), this individualistic understanding was criticized:

> The expression "the car needs petrol" is of course not an indication that the car has a feeling or an "inner motivational state." The meaning of the word "need" is that the car cannot do what *we* want it to do unless it gets some petrol. There is no reason to believe that the meaning of "need" is different when applied to human beings or to information." (Needs are thus always connected to the implicit or explicit goals of actors).

A teacher may know what information/knowledge the students need (e.g., a mathematical formula in order to calculate the movement of an object. Such a need is not primarily an internal realization or feeling by the students, but something they are taught. Patrick Wilson (1977, p. 44) made the same point:

> […] what information is relevant to a particular concern is a public, not a private, matter, with public standards of criticism. If I know your concerns, I have a basis for recognizing not merely what you are likely to want to know, but what you ought to know.

The term "information need" itself seems to be due to the individualistic failure. It may be better so say that users need to be informed by the best

research, thus removing the problem from psychology to science studies and philosophy.

Hjørland (1997, p. 172) argued that to have an "information need" is equivalent to say "documents or information of relevance to N can be found" (if no such documents exists, we should rather speak of a need for research, which can provide such documents). In other words, to specify a person's information needs involves the same process as determine relevance criteria of documents for that person. This seems obvious, but in information science, the literatures on relevance and on information needs have largely been separated. Again, researchers in information science have not considered their own information needs, and that of their colleagues. Hjørland's view (see e.g. 2024), for example, is that we in information science and KO need to know more about the philosophy of science (such as Kuhn's paradigm theory and social epistemology), other researchers may have different views. In order to develop KO as a field of study, this is not a private matter; it is a matter for the KO community. It is simply necessary to be involved in the subject matter to be able developing a position concerning a person's (or a discipline's) information need. Individuals may choose if they would like to contribute to what a discipline needs, or whether not to do so. If they choose to do so, the disciplines needs becomes their individual need.

7 Concept

Dahlberg (1978, p. 142; abstract) wrote:

> The totality of all the verifiable and necessary statements on a referent may be summarized and/or synthesized by a term which will then represent a concept in any communication process. A concept is thus regarded as a knowledge unit, and the statements about its referent are found to be the knowledge elements, also known as the characteristics, of the given concept.

In contrast to the domain analytic view, Dahlberg's definition does not open up for conflicting views on established knowledge in different theories/paradigms, and thus for conflicting concepts. Compare Kuhn (1962)

who compared concepts in the geocentric model and the heliocentric model (represented by Ptolemaic astronomers) and the heliocentric model (represented by Copernican astronomers):

a) in paradigm one, Ptolemaic astronomers might learn the concepts "star" and "planet" by having the Sun, the Moon, and Mars pointed out as instances of the concept "planet" and some fixed stars as instances of the concept "star";
b) in paradigm two, Copernicans might learn the concepts "star", "planet" and "satellites" by having Mars and Jupiter pointed out as instances of the concept "planet", the Moon as an instance of the concept "satellite" and the Sun and some fixed stars as instances of the concept "star".

Thus, the concepts "star", "planet" and "satellite" got a new meaning and astronomy got a new classification of celestial bodies.

Based on this philosophy, Hjørland (2009, p. 1522-3) defined:

> Concepts are dynamically constructed and collectively negotiated meanings that classify the world according to interests and theories. Concepts and their development cannot be understood in isolation from the interests and theories that motivated their construction, and, in general, we should expect competing conceptions and concepts to be at play in all domains at all times.

This implies that classification researchers need to relate the interests and theories, which their work serves.

8 Subject (of documents)

Users search for documents about a certain subject, and information specialists index and classify documents according to their subjects. The term "subject" is often considered a primitive term, which is not in need of a theoretical understanding. This is wrong, however, and indexers/classifiers often index or classify the same document differently. Such a lack is known as the measure called "inter indexer consistency"[8] which may be due to many factors, including mistakes or because the indexer is influenced by a certain classification system

or thesaurus. Hjørland (2017b; p. 24) found that the most important theoretical issue in determining the meaning of "subject" is implied by the two opposed theories of indexing: the "content oriented" versus the "request oriented" view. The content oriented view is represented by the so-called 20% rule used by the *Library of Congress*. According to this rule, at least 20% of any given document shall be about the subject indicated by the subject label. The request-oriented view has a very different understanding of the term "subject". It does not understand the subject of a document as something inherent in a document. It is not something that a document "has," but it is something attributed to documents in order to make them visible in relation to a particular needs or tasks. This mean, that the subject description of a given document may vary according to different user groups, or, more realistic, according to the purpose and the policy of the library or information system, which provides the indexing/subject description. (Remark the close relation to Kwon's statement in Section 4, that information informs relative to a question).

Based on such considerations, Hjørland (2017b, p. 60) defined subjects as the informative potentials of documents. If a book about Napoleon, for example, is acquired by a library devoted to feminist research, we should expect that this be done because the book is supposed to contain information about women in relation to Napoleon and his time. To increase findability to the target audience, its indexing should reflect what information of relevance to the feminist scholars it contributes. Request oriented indexing (or policy oriented indexing) also describe the contents, but does this from a certain perspective. If a small part of the book is found important for the target group, this small part is indexed although this is against the 20% rule.

9 Conclusion

The underlying issue in the examples is that information science and KO need to consider issues in philosophy of science. Hopefully, the examples have demonstrated how a theoretical stance, such as domain-analysis makes an important difference for how basic terms are understood and how mediating practices are affected. Philosophy of science is not an easy subject, but the

domain analytic point of view emphasis the study of knowledge domains, and generalized knowledge about domains, cannot ignore this field.

What, then, is the prospect of the domain analytic view? Well, it seems linked to the prospect for information science with KO. In order to understand the position better, one needs to consider whether better alternatives exist. The alternatives today seems almost absent in discussions, but just being a growing number of ad hoc fragments without a tendency to clear advancement of our discipline, which is not a fruitful alternative. Information science with KO must have a theoretical basis, and the best theoretical basis is the one based on the best arguments. This article has provided arguments for the superiority of the domain analytic view.

References

BROUGHTON, Vanda. Henry Evelyn Bliss. *In*: HJØRLAND, Birger; GNOLI, Claudio (ed.). **ISKO Encyclopedia of Knowledge Organization**. Toronto: International Society for Knowledge Organization, 2020.

BUCKLAND, Michael K. **Information and information systems**. Westport: Greenwood Press, 1991.

CASE, Donald O.; GIVEN, Lisa M. **Looking for information**: a survey of research on information seeking, needs, and behavior. 4th ed. Bingley: Emerald, 2016.

CSISZAR, Alex. Bibliography as anthropometry: dreaming scientific order at the fin de siècle. **Library Trends**, Baltimore, v. 62, n. 2, p. 442-55, 2013. Available at: https://doi.org/10.1353/lib.2013.0041. Access: Aug. 10th 2024.

DAHLBERG, Ingetraut. A referent-oriented, analytical concept theory for INTERCONCEPT. **Knowledge Organization**, Baden-Baden, v. 5, n. 3, p. 142-51, 1978. Available at: https://doi.org/10.5771/0943-7444-1978-3-142. Access: Aug. 10th 2024.

FJELDSÅ, Jon. Avian Classification in Flux. *In*: DEL HOYO, Josep; ELLIOTT, Andrew; CHRISTIE, David A (ed.). **Handbook of the birds of the world**. Barcelona: Lynx, 2013. v. 17, p.77-146, 493-501.

FLAMMER, August. Towards a theory of question asking. **Psychological Research**, New York, v.43, n. 4, p. 407-420, 1981. Available at: https://doi.org/10.1007/BF00309225. Access: Aug. 10th 2024.

FLECK, Ludwik. **Genesis and development of a scientific fact**. Translation of: Entstehung und Entwicklung einer wissenschaftlichen Tatsache (1935). Chicago: University of Chicago Press, [1935] 1979.

FUGMANN, Robert. On the role of subjectivity in establishing, using, operating and evaluating information retrieval systems. Treatise II on retrieval on retrieval system theory. **Information Storage and Retrieval**, Amsterdam, v. 9, n. 7, p. 353-372, 1973. Available at: https://doi.org/10.1016/0020-0271(73)90049-1. Access: Aug. 10th 2024.

FURNER, Jonathan. Truth, relevance, and justice: towards a veritistic turn for KO. **Advances in Knowledge Organization**, Baden-Baden, v. 16, p. 468-474, 2018.

GORICHANAZ, Tim. Questioning and understanding in the library: a philosophy of technology perspective. **Education for Information**, Clifton, v. 35, n. 4, p. 399-418, 2019. Available at: https://doi.org/10.3233/EFI-180230. Access: Aug. 10th 2024.

HJØRLAND, Birger. **Information seeking and subject representation: an activity-theoretical approach to information science**. Westport: Greenwood Press, 1997.

HJØRLAND, Birger. Concept theory. J**ournal of the American Society fo Information Science and Technology**, New Jersey, v. 60, n. 8, p. 1519-1536, 2009. Available at: https://doi.org/10.1002/asi.21082. Access: Aug. 10th 2024.

HJØRLAND, Birger. The foundation of the concept of relevance. **Journal of the American Society for Information Science and Technology**, New Jersey, v. 61, n. 2, p. 217-237, 2010. Available at: https://doi.org/10.1002/asi.21261. Access: Aug. 10th 2024.

HJØRLAND, Birger. A rejoinder to Nicolaisen's refutation of Hjørland's relevance definition. **Information Research**, Boras, v. 22, n. 1, 2017a.

HJØRLAND, Birger. Subject (of Documents). **Knowledge Organization**, Baden-Baden, v. 44, n. 1, p. 55-64, 2017b. Available at: https://doi.org/10.5771/0943-7444-2017-1. Access: Aug. 10th 2024.

HJØRLAND, Birger. Information. **Knowledge Organization**, Baden-Baden, v. 50, n. 1, p. 47-78, 2023. Available at: https://doi.org/10.5771/0943-7444-2023-1-47. Access: Aug. 10th 2024.

HJØRLAND, Birger. Social epistemology. *In*: **Knowledge Organization**, Baden-Baden, v. 51, n. 3, p. 187-202, 2024. Available at: https://doi.org/10.5771/0943-7444-2024-3-187. Access: Oct. 7th 2024.

HJØRLAND, Birger; CHRISTENSEN, Frank Sejer. Work tasks and socio-cognitive relevance: a specific example. **Journal of the American Society for Information Science and Technology**, New Jersey, v. 53, n. 11, p. 960-965, 2002. Available at: https:// https://doi.org/10.1002/asi.10132. Access: Aug. 10th 2024.

KUHN, Thomas S. The **structure of scientific revolutions**. Chicago: University of Chicago Press, 1962.

KWON, Hannah. **On the social epistemological nature of questions: a comparison of knowledge domain's question formulations on the topic "Memory"**. New Brunswick: Rutgers University, 2016. Available at: https://rucore.libraries.rutgers.edu/rutgers-lib/50032/PDF/1/. Access: Oct. 7th 2024.

LANCASTER, Frederick Wilfrid. **Information retrieval systems characteristics**: testing and evaluation. New York: Wiley, 1968.

NICOLAISEN, Jeppe. **The problem of probability**: an examination and refutation of Hjørland's relevance equation. **Information Research**, Boras, v. 22, n. 1, 2017.

SARACEVIC, Tefko. Relevance: a review of and a framework for the thinking on the notion in information science. **Journal of the American Society for Information Science**, New Jersey, v. 26, n. 6, p. 321-343, 1975. Available at: https://doi.org/10.1002/asi.4630260604. Access: Aug. 10th 2024.

SARACEVIC, Tefko. Relevance Reconsidered. *In*: CONFERENCE ON CONCEPTIONS OF LIBRARY AND INFORMATION SCIENCE INFORMATION SCIENCE (CoLIS 2), 2., 1996, Copenhagen . **Proceedings** […].Copenhagen: Royal School of Librarianship, 1996. p. 201-218.

SARACEVIC, Tefko. Research on relevance in information science: a historical perspective. *In*: Carbo, Toni; Bellardo Hahn, T. (ed.). **Proceedings of the American Society for Information Science and Technology (ASIS&T) 2012**. Preconference on the History of ASIS&T and Information Science and Technology. Medford: Information Today, 2012. p. 49-60.

SHERA, Jesse H. Classification as the basis of bibliographic organization. *In*: SHERA, Jesse H.; EGAN, Margaret E. (ed). **Bibliographic organization**: papers presented before the Fifteenth Annual Conference of the Graduate Library School July 24-29, 1950. Chicago: University of Chicago Press, 1951. p. 72-93.

SOERGEL, Dagobert. Is user satisfaction a Hobgoblin? **Journal of the American Society for Information Science**, New Jersey, v. 27, n. 4, p. 256-259, 1976. Available at: https://doi.org/10.1002/asi.4630270411. Access: Aug. 10th 2024.

TAYLOR, Robert S. Question-negotiation and information seeking in libraries. **College & Research Libraries**, Chicago, v. 29, n. 3, p. 178-194, 1968. Available at: https://doi.org/10.5860/crl.76.3.251. Access: Aug. 10th 2024.

VAN RIJSBERGEN, Cornellis J.; LALMAS, Mounia. Information calculus for information retrieval. **Journal of the American Society for Information Science**, New Jersey, v. 47, n. 5, p. 385-98, 1996. Available at: https://doi.org/10.1002/(SICI)1097-4571(199605)47:5<385::AID-ASI6>3.0.CO;2-S. Access: Aug. 10th 2024.

WHITE, Howard D. Relevance in Theory. *In*: MCDONALD, John D.; LEVINE-CLARK, Michael (ed.). **Encyclopedia of Library and Information Sciences**. 4th. ed. Boca Raton: CRC Press, 2017. p. 3926-3939.

WILSON, Patrick. **Public knowledge, private ignorance**. Westport: Greenwood Press, 1977.

ZHANG, Lei. A Tale of indexing (in)consistency. **The Indexer: The International. Journal of Indexing**, Liverpool, v. 38, n. 2; p. 171-183, 2020. Available at: https://doi.org/10.3828/indexer.2020.16. Access: Aug. 10th 2024.

Alguns conceitos-chave da análise de domínio

Resumo: O presente trabalho parte do pressuposto que os sistemas de informação devem fornecer informações confiáveis e dignas de confiança. Esta demanda implica uma reorientação da ciência da informação com a organização do conhecimento: por exemplo, de abordagens baseadas em estudos das preferências dos usuários para estudos de qualidade e dos valores e objetivos nos documentos que estão sendo mediados. A mediação da informação é realizada por profissionais de informação em seus sistemas e processos de organização do conhecimento. Para esse fim, a filosofia da ciência torna-se um assunto importante, com implicações para a compreensão de uma ampla gama de conceitos. Aqui, os seguintes conceitos são brevemente apresentados: “epistemologia social”, “informação”, “questão”, “relevância"” “necessidade de informação”, “conceito” e “assunto” (dos documentos).

Palavras-chave: análise de domínio; epistemologia social; necessidade de informação; informação

Received: 08/07/2024
Accepted: 08/08/2024

Authorship and responsibility statement

Conception and design of the study: Birger Hjørland
Data collection: Birger Hjørland
Data analysis and interpretation: Birger Hjørland
Writing: Birger Hjørland
Critical review of the manuscript: Birger Hjørland

How to cite

HJØRLAND, Birger. Some core concepts of domain analysis. **Em Questão**, Porto Alegre, v. 30, e-141781, 2024. DOI: https://doi.org/10.1590/1808-5245.30.141781

[1] During the history of the DDC, there has been a competition of two opposite principles, "the integrity of numbers" and "keeping pace with knowledge". The first of these principles mean that numbers in the DDC should not be changed, even if this is required to update the system to reflect current knowledge. Melvil Dewey himself was a staunch defender of the first principle, which he maintained in the early editions as the system, and he even strongly cautioned classification systems against trying to represent scientific developments (see Csiszar 2013, p. 445). Some later editors have supported the opposite view, but even today, the DDC is characterized by prioritizing stability at the expense of updated knowledge.

[2] See Broughton (2020, Section 3.3.1.1: https://www.isko.org/cyclo/bliss#3.3.1.1).

[3] Fjeldså (2013, p. 141) describe how many kinds of birds until very recently were considered blackbirds or subspecies of blackbirds: "Thus, rather than treating these blackbirds as different subspecies or as closely allied species, we can regard them as only convergently similar, as the males independently developed a black plumage, contrasting the yellow bill, as an effective means of demonstrating dominance within their territory."

[4] Saracevic (1996, p. 201) wrote: "Thus, not only information, but information characterized by its relevance became the key notion in information science. And the key headache."

[5] Comment BH: The call for more research to clarify the concept of relevance is, of course, correct, if it is wrong that it is badly understood. But when researchers such as Saracevic neglect former insights (such as "the subject knowledge view of relevance," research does not advance, but become muddled. So much is written about relevance today, that it seems to be a research goal itself to write articles about this concept in absence from the problems it is meant to solve for us.

[6] Hjørland's definition of relevance was challenged by Nicolaisen (2017) and Nicolaisen's criticism refuted by Hjørland (2017a). The most important objection to Nicolaisen is to this

sentence: "Recall that it is Hjørland who maintains that it is possible to construct an equation that can be used for assessing the relevance of anything." However, what Hjørland has claimed is a definition of relevance, not that to be able to "assessing the relevance of anything." Nicolaisen confuses the understanding and definition of the concept of relevance, with the assessment of whether a specific thing (document, knowledge claim) is relevant for a given purpose.

[7] Saracevic (2012, p. 51) wrote: "Bibliographic classifications, subject headings, and indexing languages were used for organizing information or information records for a long time, some schemes and practices going back centuries. All are based on the notion of *aboutness*. Choice of a given classification code, subject heading, or index term denotes what a document, or part thereof, is *about*. They assume but do not address searching at all. Searching is taken for granted. In other words, all deal with inputs and take outputs as a given. No attempt was ever made to define searching related to any classification or subject heading scheme, be it formally or pragmatically."

[8] "Inter indexer consistency" is discussed by Zhang (2020) and references herein.

Birger Hjørland

Some core concepts of domain analysis
Alguns conceitos-chave da análise de domínio

Em Questão
vol. 30, e-141781, 2024
Universidade Federal do Rio Grande do Sul,
ISSN: 1807-8893
ISSN-E: 1808-5245

DOI: https://doi.org/10.1590/1808-5245.30.1417