Reviews of Concepts in Knowledge Organization
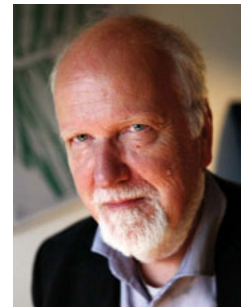
Series Editor: Birger Hjørland

Subject (of Documents) †‡

Birger Hjørland

University of Copenhagen, Royal School of Library and Information Science, South Campus Bldg 4, Room 4A-2-70, Njalsgade 76, DK-2300 Copenhagen S, Denmark, <birger.hjorland@hum.ku.dk>

Abstract: This article presents and discusses the concept "subject" or subject matter (of documents) as it has been examined in library and information science (LIS) for more than 100 years. Different theoretical positions are outlined and it is found that the most important distinction is between document-oriented views versus request-oriented views. The document-oriented view conceives subject as something inherent in documents, whereas the request-oriented view (or the policy-based view) understands subject as an attribution made to documents in order to facilitate certain uses of them. Related concepts such as concepts, aboutness, topic, isness and ofness are also briefly presented. The conclusion is that the most fruitful way of defining "subject" (of a document) is the document's informative or epistemological potentials, that is, the document's potentials of informing users and advancing the development of knowledge.

1.0 Introduction

In library and information science (LIS), documents (such as books, articles and pictures) are classified, indexed and searched by "subject" (as well as by other attributes such as author, genre and language). This makes "subject" a fundamental concept in this field (see Golub 2014 for a recent text). This use of "subject" in LIS is part of the broader use of the concept that refers to all kinds of utterances ("what is he talking about"). LIS specialists assign subject labels to documents to make them findable/retrievable. Such professionally assigned subject labels compete with other subject access points such as words from titles, abstracts and full-text, bibliographic references, user tagging, etc. Therefore, research in subject representation is not limited to professionally assigned subject labels but includes the study of all possible subject access points.

56

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

There are many ways to produce subject representations, and in general there is not always consensus about which subject should be attributed to a given document. As stated by Lancaster (2003, 21), it is important to distinguish the conceptual analysis and the translation stages in indexing and classification. In conceptual analysis, subjects are attributed to documents, and in the translation stage subject labels are assigned to documents. There tends to be great variation among indexers and classifiers in subject analysis and choice of subject labels, as measured, for example, by so-called inter-indexer consistency studies (see Saracevic 2008). To optimize subject representation and searching, we need to have a deeper understanding of the questions

– What is the criterion that a given subject should be attributed to a given document?
– What is to be understood by the statement "document A belongs to subject category X?"
– What is a subject?

This issue has been debated in the field for more than 100 years, often by using other terms such as "aboutness" or "topic" (cf., below).

One may think that the concept "subject" in this connection is self-evident and in no need for theoretical exploration. The claim of this article is, however, that it is a basic concept with different meanings and that a fruitful understanding of it is of fundamental importance for LIS. What Tredinnick (2006, 19) wrote about the concepts information, knowledge, data, document and text is equally true for subject:

> The difficulty in reaching agreement about their meaning in part derives from the kinds of research questions that are addressed, but also in part from fundamental differences in the conceptual outlooks into which they are slotted. Implicit in this is an ongoing cycle of appropriation and reappropriation of the meaning of these contested terms for particular ends.

Therefore, we have to consider the different theoretical outlooks in order to decide which outlook and thereby understanding of "subject" is most fruitful for knowledge organization.

2.0 Theoretical views

This section provides a chronological presentation of definitions or understandings of "subject" in LIS. The presentation seeks to present all significant views without guarantee of being complete (the presentation has been

difficult to produce because different researchers have mostly ignored former definitions).

2.1 Charles Ammi Cutter (1837–1903)

For Cutter, the stability of subjects depends on a social process in which their meaning is stabilized in a name or a designation. We are here presenting Cutter's view from Miksa (1983a) and Frohmann (1994).

Francis Miksa wrote: "[a subject] referred … to those intellections … that had received a name that itself represented a distinct consensus in usage" and the "systematic structure of established subjects [is] resident in the public realm" (Miksa 1983a, 69); "[s]ubjects are by their very nature locations in a classificatory structure of publicly accumulated knowledge" (Miksa 1983a, 61).

Bernd Frohmann added (1994, 112-13):

> The stability of the public realm in turn relies upon natural and objective mental structures which, with proper education, govern a natural progression from particular to general concepts. Since for Cutter, mind, society, and SKO [systems of knowledge organization] stand one behind the other, each supporting each, all manifesting the same structure, his discursive construction of subjects invites connections with discourses of mind, education, and society. The *Dewey Decimal Classification (DDC)*, by contrast, severs those connections. Melvil Dewey emphasized more than once that his system maps no structure beyond its own; there is neither a "transcendental deduction" of its categories nor any reference to Cutter's objective structure of social consensus. It is content-free: Dewey disdained any philosophical excogitation of the meaning of his class symbols, leaving the job of finding verbal equivalents to others. His innovation and the essence of the system lay in the notation. The *DDC* is a poorly semiotic system of expanding nests of ten digits, lacking any referent beyond itself. In it, a subject is wholly constituted in terms of its position in the system. The essential characteristic of a subject is a class symbol which refers only to other symbols. Its verbal equivalent is accidental, a merely pragmatic characteristic …. The conflict of interpretations over "subjects" became explicit in the battles between "bibliography" (an approach to subjects having much in common with Cutter's) and Dewey's "close classification." William Fletcher spoke for the scholarly bibliographer …. Fletcher's "subjects," like Cutter's, referred to the categories of a fantasized, stable social order, whereas Dewey's subjects were elements of a semiological system of standardized, techno-

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

57

bureaucratic administrative software for the library in its corporate, rather than high culture, incarnation.

Cutter's view on "subject" is probably wiser than most of the later understandings that dominated the twentieth century, including the understanding reflected in the International Standards Organization (ISO) standard quoted below. The early statements quoted by Frohmann indicate that subjects are somehow shaped in social processes. It also indicates that there was a conflict between Cutter and Dewey in understanding "subjects" that is reflected in their respective classification systems. When that is said, it should be added that Cutter's view seems not particularly detailed or clear. We only get a vague idea of the social nature of subjects.

2.2 S. R. Ranganathan (1892–1972)

Ranganathan provided the following definitions:

> "Subject: assumed term" (Ranganathan 1963, 27).
> "Subject: Thought-content of a document" (Ranganathan 1964, 109).
> "Subject—an organized body of ideas, whose extension and intention are likely to fall coherently within the field of interests and comfortably within the intellectual competence and the field of inevitable specialization of a normal person" (Ranganathan 1967, 82).

A related definition is given by one of Ranganathan's students (Gopinath 1976, 51): "A subject is an organized and systematized body of ideas. It may consist of one idea or a combination of several."

The first of Ranganathan's definitions (1963) seems to consider "subject" self-evident and in no need for theoretical exploration. His second definition (1964) corresponds to the "content-oriented view" (see section 2.4 below). Ranganathan's third definition (1967) as well as that of Gopinath (1976), are here taken as the point of departure and, overall, they are considered alike. Ranganathan's 1967 definition of "subject" is clearly influenced by his Colon Classification system (CC), which is an analytic-synthetic scheme based on the combination of single elements from facets to subject designations. The definition needs to be understood in the context of his other concepts such as "isolate" and "basic concept" in CC. Ranganathan's concepts are highly idiosyncratic, for example, the claim that gold cannot be a subject (but is alternatively termed "an isolate"). The concept "discipline" is substituted with "basic subject" defined (1967, 83) as a subject, that does not have isolate ideas as a component (example: mathematics).

We can see a problem with Ranganathan's concepts if we consider a simple sentence. In the Dewey system (*DDC*) it is stated "No other feature of the DDC is more basic than this, that it scatters subjects by discipline (Dewey 1979, xxxi)." This makes sense, and "subject" as well as "discipline" are here used in a way that is not specific for *DDC*, but can be applied generally. This is not so with Ranganathan's concepts, which can only be understood in relation to CC.

If we consider the 1967 definition with the definitions presented in the rest of this article, we can see that it provides no guidance in itself for subject analysis. It does not address the problems, for example, raised by Wilson (section 2.3) or by issues discussed in section 2.4 and 2.5. That a subject is organized seems just to refer to how subjects are analyzed in CC, where subjects are organized and combined of single elements from facets. This is the reason why the organized or combined nature of subjects is emphasized. It seems unacceptable that Ranganathan defines the concept of "subject" in a way that favors his own system. A scientific concept like "subject" should make it possible to compare different ways of establishing access to information. Whether we speak of, for example, of enumerated systems, pre-coordinative systems, faceted systems or post-coordinative systems, whether subjects are organized or not, should not be a part of the definition of "subject" (but when "subject" has been defined, its degree of organization may be examined in specific cases). Ranganathan's definition also contains the pragmatic demand that a subject should be determined in a way that suits a normal person's competency or specialization. Again, we see a strange kind of mixing a general understanding of a concept with demands put by a specific system. One thing is what the concept "subject" represents; quite another issue is how to provide subject descriptions that fulfill demands such as precision and recall. Because Ranganathan's (1967) definition is too closely related to his CC, comparative studies of different kinds of systems are made difficult by using it.

This aspect of the theory was criticized by Metcalfe (1973, 318). Metcalfe's skepticism regarding Ranganathan's theory is formulated in harsh words (op. cit., 317): "This pseudo-science imposed itself on British disciples from about 1950 on." Although this voice is contrary to Ranganathan's generally high prestige in LIS and partly dismissed by Drake (1960), it seems important that Ranganathan's theoretical assumptions be carefully examined and not taken for granted (see also Hjørland 2013). Ranganathan's concept of subject has been further presented by Dutta (2015), Dutta and Dutta (2013) and Dutta, Majumder and Sen (2013). These articles are, however, mostly summaries of other authors' papers.

58

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

Based on these arguments we may conclude that Ranganathan's definition of the concept of "subject" is not suited for general scientific use. Like the definition of "subject" given by the ISO-standard for topic maps (see Section 2.7), Ranganathan''s definition may be useful within his own closed system. The purpose of a scientific and scholarly field is, however, to examine the relative fruitfulness of systems such as topic maps and CC. For such purposes, another understanding of the concept of "subject" seems to be necessary.

2.3. Patrick Wilson (1927–2003)

In his book, Wilson (1968) examined—in particular by thought experiments—the suitability of different methods of determining the subject of a document. The methods were:

– To identify the author's purpose for writing the document.
– To weigh the relative dominance and subordination of different elements in the picture, which the reading imposes on the reader.
– To group or count the document's use of concepts and references.
– To construe a set of rules for selecting elements deemed necessary (as opposed to unnecessary) for the understanding of the work as a whole.

Wilson demonstrated that each of these methods is insufficient to determine the subject of a document. He is led to conclude that (Wilson 1968, 89) "The notion of the subject of a writing is indeterminate," or about what users may expect to find using a particular position in a library classification system (Wilson 1968, 92): "For nothing definite can be expected of the things found at any given position."

In connection with the last quote, Wilson (1968, 92) adds an interesting footnote in which he writes:

> For example, I know more or less clearly what hostility is, that is, the word 'hostility' has a fairly sharp meaning for me, but far from a perfectly sharp and precise meaning. Now if I were to supply myself with an exact defined concept, got by explication of my imprecise notion, I might find that I could never use the new concept in describing any actual piece of writing; the concept might be too sharp ever to find application. There would be instances of hostility (in the new sense) that I could recognize, but no instances of writings on hostility that I could recognize, for no one would have written on hostility (as I now would understand it). If people write on what

are for them ill-defined phenomena, a correct description of their subjects must reflect the ill-definedness.

Hjørland (1992) discussed Wilson's concept of subject and found that it is problematic to give up the precise understanding of such a basic concept in LIS. Wilson's arguments led him to an agnostic position, which Hjørland found unacceptable and unnecessary. Concerning the authors' use of ambiguous terms, the role of the subject analysis is to determine which documents would be fruitful for users to identify whether or not the documents use one or another term or whether a given term in a document is used in one or another meaning. The information specialists provide an interpretation and a description (for example based on a controlled vocabulary) which classify the literature in a way that users may learn to use to identify the terms or classes that with high probability refer to the needed documents. Relevant concepts and distinctions in classification systems and controlled vocabularies may be fruitful even if applied to documents with ambiguous terminology. The problem is not whether there is a precise match between the documents' and the information specialist's concepts, but whether the subject representation makes distinctions that are relevant for the users.

2.4 "Content-oriented" versus "request-oriented" views

In this section, two kinds of indexing principles will be presented that illuminate a core theoretical issue related to the concept "subject." Traditional indexing has been content or document oriented. An example is the "20%" rule used by, for example, the Library of Congress. According to this rule, at least twenty percent of any given document shall be about the subject indicated by the subject label (Library of Congress 2008, sheet H 180):

> Assign headings only for topics that comprise at least 20% of the work.
> In the case of a work containing separate parts, for example, a narrative text plus an extensive bibliography or a section of maps (cf. H 1865), or a book with accompanying materials, such as a computer disc, assign separate headings for the individual parts or materials if they constitute at least 20% of the item and are judged to be significant.

The alternative principle is request-oriented indexing in which the anticipated request from users is influencing how documents are being indexed. The indexer asks himself or herself: "Under which descriptors should this entity [document] be found?" and "think of all the possible

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

59

queries and decide for which ones the entity at hand is relevant" (Soergel 1985, 230; see also Soergel 1974, Chapter F1, 356).

Request-oriented indexing may be indexing that is targeted towards a particular audience or user group. A library for feminist studies may, for example, index documents differently compared to a historical library: If a feminist library buys a book of, say, Napoleon, it must be assumed that it does so because the book in some way is relevant from a feminist perspective (i.e., says something about women at the time of Napoleon). For the user of the catalog, it is important that this purpose and perspective be expressed in the subject representation of the book, in order to enable users to find books about women at that time. In other words, the purpose or perspective of a specific collection should ideally be reflected in its classification or indexing (which, of course, is contrary to economic considerations to standardize subject representation and reuse the work done by other libraries, there are thus contradicting interests at play). It is probably best to understand request-oriented indexing as policy-based indexing: as indexing done according to some ideals and reflecting the purpose of the library or database for which it is done. In this way, it is not necessarily a kind of indexing based on empirical user requests, but only those anticipated requests that are considered within the purpose of the library or database to answer (only if empirical data about use or users are applied, should indexing be regarded as user-based).

It is interesting to consider that mainstream automatic indexing is not purely document oriented because the frequency of terms in a given collection is taken into account. Terms that are used in many documents have a low discriminatory power and are therefore assigned a lower weight. In this way automatic indexing is less document oriented and more contextual compared with, for example, the use of the twenty percent rule. Still, of course, this principle of automatic indexing does not fulfill the demands of "request-oriented indexing"/ "policy-oriented indexing."

The content-oriented view considers "subject" to be something inherent in documents. The request-oriented view, on the other hand, considers "subject" to be something attributed to documents by somebody in order to facilitate certain uses of the documents. The problem of whether the subject is something inherent in the document (and determined "objectively") or is context dependent (and determined "subjectively") is related to the philosophical subject–object problem to which we now turn.

2.5 Issues of subjectivity and objectivity

There exists an ideal, often implicit, that there is one right way to provide a subject representation for a given docu-

ment. The formerly mentioned inter-indexer consistency studies is then an example of an attempt to measure the subjectivity in subject representation based on the assumption that the majority of indexers are closer to the truth compared to the outliers. However, as pointed out by Cooper (1969), indexing may be consistently wrong. Indexers may be guided, for example, by the same bad principles or assumptions and in that case their indexing will be consistently bad. Therefore, studies of inter-indexer consistency may not necessarily provide a basis for improving indexing quality. The implication is that we can only determine the quality of subject representation from the standpoint of a theory of what good indexing and classification should be like. If we take the request-oriented view as the point of departure, the subject representation should only be consistent in relation to the same anticipated requests or the same policy framework. In other words, subject representation should be based on intersubjectively stated goals, values and policies. They should not as an ideal be objective. In a way, the subjectivity of indexers should be an ideal (but not any form of subjectivity, of course, just a subjectivity developed to consider a specific perspective).

"Subject" may also mean the knowing subject (person) who retrieves documents that answer questions for him. In general, these two meanings are separated in information science, although, as we saw above, different persons may provide different subject representations, even as an ideal. In a recent monograph (Day 2014) these two meanings of "subject" are combined. The main point in Day's book is that indexes in a certain theoretical perspective (Day 2014, 37) "have more than simply a retrieval function; they do not only act as affordances and means for the fulfillment of 'information needs,' but for the creation of such, and the creation of documentary-mediated persons and selves, as well"—a point of view that may seem somewhat exaggerated.

2.6 The subject knowledge view

The subject knowledge view of subjects emphasizes the role of domain specific knowledge in relation to both subject representation in practice and theoretical issues concerning the nature of "subject." It may also be called "the domain analytic view" or "the epistemological view" because it understands subject knowledge as formed by different theories, which in the end are connected with epistemological assumptions. Rowley and Hartley wrote (2008, 109)

In order to achieve good consistent indexing, the indexer must have a through appreciation of the structure of the subject and the nature of the contribu-

60

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

tion that the document is making to the advancement of knowledge within a particular discipline.

This is an important statement (which unfortunately has not been further developed by the authors). It clearly expresses that subject representation aims at supporting advancement of knowledge in different domains and that subject knowledge is a precondition for doing so. This statement is in accordance with how Hjørland (1992, 185) defined subjects as the epistemological potentials of documents (or, synonymously, as the informative potentials of documents). This definition also implies that subject representations aim at supporting advancement of knowledge in different domains and that subject knowledge is a precondition for doing so. Hjørland's definition contains the additional layer that different "paradigms" entail different subject representations. Therefore, the question of subject representation is closely linked to the question of which paradigms should be supported. In other words, subject representations cannot be regarded as neutral expressions. On the contrary, the activity of assigning a subject label to a given document represents a kind of power, (cf., Olson 2002) which aims at facilitating certain uses of that document at the expenses of other uses.

Let us consider a concrete example. Fisher (1921), as a part of a series, published the article "Studies in Crop Variation." As indicated by the title, the subject is "crop variation." Retrospectively, however, this title and subject attribution is considered poor (Salsburg 2001, 43):

> Seldom in the history of science has a set of titles [*Studies in Crop Variation*] been such a poor description of the importance of the material they contain. In these papers, Fisher developed original tool for the analysis of data, derived the mathematical foundations of those tools, described their extensions into other fields, and applied them to the "muck" he found at Rothamsted. These papers show a brilliant originality and are filled with fascinating implications that kept theoreticians busy for the rest of the twentieth century, and will probably continue to inspire work in the years that follow.

Of course, Fisher (1921) is (also) about "crop variation" and should be indexed as such in indexes within agriculture. However, as the quote says, this article has had a much broader and deeper importance in the field of statistical probability where two of its main subjects are "experimental design" and "sampling." If the purpose of subject representation is to support future use of documents, then these last mentioned subject labels are far more important than that indicated by the title.

The bibliometrican Henry Small published an important paper "Cited Documents as Concept Symbols" (Small 1978) in which he found that highly cited papers tend to be cited for the same reasons and that these reasons are often represented in the citing documents as "concept symbols." For example, we may assume that most of the papers citing Fisher (1921) use, for example, "experimental design" as a concept symbol at the place of the reference in the text. Bibliometric methods may therefore be used automatically or semiautomatically to determine the subject of documents in a way that is in agreement with the subject knowledge view (cf., Schneider and Borlund 2004). Of course, this technique cannot be applied to assign subject labels to new documents, only retrospectively and only to (highly) cited documents. Whether or not we may apply this method in practice, the example provides a deep insight to the dynamic nature of "subjects." It demonstrates that the subject of a document is not independent of evaluation of the potentials of that document.

2.7 Other views and definitions

In the ISO-standard for topic maps, the concept of subject is defined this way (ISO/IEC 13250 2002, 4): "Anything whatsoever, regardless of whether it exists or has any other specific characteristics, about which anything whatsoever may be asserted by any means whatsoever." This definition may work well with the closed system of concepts provided by the topic maps standard. In broader contexts, however, it is not fruitful because it does not contain any specification on how to determine the subject of a given document. If different methods of subject analysis imply different results, which of these results should then be preferred? Different persons may have different opinions about what the subject of a specific document is. The theoretical understanding of the concept of "subject" should be helpful for deciding principles of subject analysis. It is not helpful just to say "subject" is "anything whatsoever."

3.0 Related concepts

3.1 Words versus concepts versus subjects

A proposal for the differentiation between concept indexing and subject indexing was given by Bernier (1980). In his opinion, subject indexes are different from, and can be contrasted with, indexes to concepts and words. Subjects are what authors are working and reporting on. A document can have the subject of "chromatography" if this is what the author wishes to inform about. Papers using chromatography as a research method or discussing it in a subsection do not have chromatography as sub-

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

61

jects. Indexers can easily drift into indexing concepts and words rather than subjects, but this is not good indexing.

Bernier does not, however, differentiate authors' subjects from those of the information seekers. A user may want a document for other reasons that its author intended. From the point of view of information systems, the subject of a document is related to the questions that the document can answer for the users (cf. the distinction between a content-oriented and a request-oriented approach presented above).

This distinction between words, concepts and subjects is often confused. If "subject" is defined differently from words and concepts, it follows that its statistical distribution may also be different. Hjørland and Nicolaisen (2005) in their analysis of the concept of "subject" in relation to Bradford's law of scattering made this distinction:

– Lexical scattering is the scattering of words in texts and in collections of texts.
– Semantic scattering is the scattering of concepts in texts and in collections of texts.
– Subject scattering is the scattering of items useful to a given task or problem.

This examples demonstrates that the concept of subject has wide-ranging implications not just for subject representation but also for bibliometrics and LIS in general.

3.2 Aboutness

Aboutness is a concept used in LIS, linguistics, philosophy of language and philosophy of mind. In the philosophy of mind, it has been often considered synonymous with intentionality (cf., Siewert 2016); in the philosophy of logic and language it is understood as the way a piece of text relates to a subject matter or topic (cf., Demolombe and Jones 1999; Yablo 2014).

Robert A. Fairthorne (1969) is credited with coining the term "aboutness" in LIS, which became popular in LIS in the late 1970s, perhaps due to arguments put forward by William John Hutchins (1975; 1977; 1978). Hutchins argued that "aboutness" was to be preferred to "subject" because it removed some epistemological problems (e.g., that different people may attribute different subjects to the same document). Hjørland (1992; 1997) argued, however, that the same epistemological problems were also present in Hutchins' proposal (different people may also attribute different "aboutness" to the same document). Because the same problems are connected with "aboutness," the reason to introduce this term as a substitute for "subject" is unsupported. By implication, "aboutness" and "subject" should be considered synonymous in LIS.

Tredinnick (2006) throughout the book considers the attribution of "aboutness" to documents to be a problematic activity in LIS ("subject" is not discussed). He wrote (Tredinnick 2006, 138):

Any isolation of the *aboutness* of texts therefore involved an act of interpretation that seeks to limit the signifying value of the text, without any particular claim to authority or authenticity. In other words, what information means also becomes a matter of the socio-cultural values that we bring to it, what Eco (1976) calls the cultural codes within which signification occurs, and these values are neither neutral in the way we might assume, nor absolute. The identification of the aboutness of information imposes certain privileged perspectives on text. It happens that these perspectives can be mapped against sociocultural norms or particular discursive communities, such as the humanist outlook that influenced librarianship and the positivism of information science. This is a problem for the information profession, which largely occupies itself by isolating in various ways the *aboutness* of texts.

If I understand this quote correctly, it says that the determination of aboutness involves socio-cultural values (and this covers "subject" as well). It is difficult to understand, however, that this act in and of itself is considered a problem; it should only be considered a problem if epistemological and socio-cultural values are ignored.

3.3 Topic

"Topic" is a term often used synonymously with "subject" and "aboutness." Examples are Jarneving (2005, 252), who wrote, "title words have a high topicality;" Xu and Yin (2008, 202) wrote: "Topicality measures the 'aboutness' of a document to the topic area suggested by a query" and Janes (1994, 161) wrote "Topicality, the relation of a document to the topic of a user's query."

Based on how the term "topic" is used in the literature of LIS, it is here concluded that it should be considered a synonym for "subject."

3.4 Isness

"Isness" is a concept that has been suggested to cover terms for indexing that are considered to be beyond proper subject terms. The International Federation of Library Associations and Institutions wrote (IFLA 2010, 10):

The FRSAR Working Group is aware that some controlled vocabularies provide terminology to express

62

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

other aspects of works in addition to subject (such as form, genre, and target audience of resources). While very important and the focus of many user queries, these aspects describe isness or what class the work belongs to based on form or genre (e.g., novel, play, poem, essay, biography, symphony, concerto, sonata, map, drawing, painting, photograph, etc.) rather than what the work is about.

"Isness" thus expresses what something is as opposed to what it is about. It is however a rather seldomly used term in LIS.

3.5 Ofness

In picture indexing, the term "ofness" is sometimes used to refer to objects or events in the picture (IFLA 2010, 11):

> Those LIS authors who have focused on the subjects of visual resources, such as artworks and photographs, have often been concerned with how to distinguish between the "aboutness" and the "ofness" (both specific and generic depiction or representation) of such works (Shatford 1986). In this sense, "aboutness" has a narrower meaning than that used above. A painting of a sunset over San Francisco, for instance, might be analyzed as being (generically) "of" sunsets and (specifically) "of" San Francisco, but also "about" the passage of time.

Shatford's analysis was inspired by Panofsky (1939), who identified three levels of meaning in works of art. At the first, or pre-iconographic, level, subject matter was designated as factual ("ofness") or expressional ("aboutness"), and based on the objects and events in an image as it could be interpreted through everyday experience. At the second, or iconographic, level, interpretation requires some cultural knowledge of themes and concepts (not "a sailor" but "Ulysses"). The third or iconological level requires interpretation at a sophisticated level using world and cultural knowledge plus a deeper understanding of the history and background of the work (see further in: Baca and Harpring (2000), Krause (1988) and Shatford (1986)).

4.0 Conclusion

The concept "subject" has a long history in LIS, but the different meanings have seldom been compared and examined. The main conclusions of this article are:

– Any approach to subject representation is connected to a certain understanding of "subject," which is often implicit.

– Different definitions or implicit views of "subject" are connected to different approaches and paradigms in information science. The concept "subject" cannot be properly understood or developed without considering basic theoretical issues in LIS.
– The activity of assigning a subject label to a given document aims at facilitating certain uses of that document at the expenses of other uses. This activity is done by somebody or by an algorithm based on his or her (or the programmer's) knowledge, theories, working conditions, etc.
– Any given document has an unlimited range of possible uses or potentials. The aim of subject analysis is to identify the most important potentials in order to facilitate the identification of documents that supports important human activities. The subjects of a document are its informative or epistemological potentials, that is its potential of informing users and advance the development of knowledge.

References

Baca, Murtha and Harpring, Patricia (Eds.). 2009. *Categories for the Description of Works of Art (CDWA)*. Los Angeles, CA: The J. Paul Getty Trust and College Art Association, Getty Research Institute. http://www.getty.edu/research/conducting%5Fresearch/standards/cdwa/index.html

Bernier, Charles L. 1980. "Subject Indexes." In *Encyclopedia of Library and Information Science: Volume 29,* edited by Allen Kent, Harold Lancour and Jay E. Daily. New York, NY: Marcel Dekker, Inc., 191-205.

Cooper, William S. 1969. "Is Interindexer Consistency a Hobgoblin?" *American Documentation* 20: 268-78.

Day, Ronald E. 2014. *Indexing It All: The Subject in the Age of Documentation, Information, and Data*. Cambridge, MA: The MIT Press.

Demolombe, Robert and Andrew J. I. Jones.1999. "On Sentences of the Kind Sentence 'P' Is about Topic 't.'" In *Logic, Language and Reasoning. Essays in Honor of Dov Gabbay,* ed. H-J. Ohlbach and U. Reyle. Dordrecht: Kluwer, 125-44. https://www.irit.fr/~Robert.Demolombe/publications/1996/gabbay96.pdf

Dewey, Melvil. 1979. *Dewey Decimal Classification and Relative Index*, 19th ed., v. 1. Albany, NJ: Forest Press.

Drake, Cyril Lewis. 1960. "What is a Subject?" *Australian Library Journal* 9: 34-41.

Dutta, Bidyarthi. 2015. "Ranganathan's Elucidation of 'Subject' in the Light of 'Infinity (∞)'." *Annals of Library and Information Studies* 62: 255-64. http://nopr.niscair.res.in/bitstream/123456789/33720/1/ALIS%2062(4)%20255-264.pdf

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

63

Dutta, Bidyarthi and Chaitali. Dutta. 2013. "Concept of 'Subject' in Library and Information Science from a New Angle." *Annals of Library and Information Studies* 60 no. 2: 78-87.

Dutta, Bidyarthi, Krishnapada Majumder and B. K. Sen. 2013. "In Search of Dimensions of Subject from the Standpoint of Ranganathan." *Annals of Library and Information Studies* 60, no. 1: 51-5.

Eco, Umberto. 1976. *A Theory of Semiotics*. Bloomington: Indiana University Press.

Fairthorne, Robert A. 1969. "Content Analysis, Specification and Control." *Annual Review of Information Science and Technology* 4: 73-109.

Fisher, Ronald Aylmer. 1921. "Studies in Crop Variation. I. An examination of the Yield of Dressed Grain from Broadbalk." *Journal of Agricultural Science* 11: 107–35. doi: 10.1017/S0021859600003750

Frohmann, Bernd. 1994. "The Social Construction of Knowledge Organization: The Case of Melvil Dewey." In *Knowledge Organization and Quality Management: Proceedings of the Third International ISKO Conference 20-24 June 1994 Copenhagen, Denmark*, ed. Hanne Albrechtsen and Susanne Oernager. Advances in Knowledge Organization 4. Frankfurt/Main: Indeks Verlag, 109-17.

Golub, Koraljka. 2014. *Subject Access to Information: An Interdisciplinary Approach*. Santa Barbara, CA: Libraries Unlimited.

Gopinath, Malur Aji. 1976. "Colon Classification." In *Classification in the 1970s: A Second Look* (rev. ed.), ed. Arthur Maltby. London: Clive Bingly, 51-80.

Hjørland, Birger. 1992. "The Concept of 'Subject' in Information Science." *Journal of Documentation* 48: 172-200.

Hjørland, Birger. 1997. *Information Seeking and Subject Representation: An Activity-Theoretical Approach to Information Science*. Westport & London: Greenwood Press.

Hjørland, Birger. 2013. "Facet Analysis: The Logical Approach to Knowledge Organization." *Information Processing and Management* 49: 545-57.

Hjørland, Birger and Jeppe Nicolaisen. 2005. "Bradford's Law of Scattering: Ambiguities in the Concept of 'Subject'." In *CoLIS 2005, Proceedings of the 5th International Conference on Conceptions of Library and Information Science*, ed. F. Crestani and I. Ruthven. Berlin: Springer-Verlag, 96-106.

Hutchins, W. John. 1975. *Languages of Indexing and Classification: A Linguistic Study of Structures and Functions*. London: Peter Peregrinus.

Hutchins, W. John. 1977. "On the Problem of 'Aboutness' in Document Analysis." *Journal of Informatics* 1: 17-35.

Hutchings, W. John. 1978. "The Concept of 'Aboutness' in Subject Indexing." *Aslib Proceedings* 30: 172-81.

IFLA. 2010. *Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model*. Berlin: De Gruyter. http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf

ISO/IEC 13250 *Topic Maps. Information Technology. Document Description and Processing Languages*, 2nd ed. Geneva, 19 May 2002. http://xml.coverpages.org/TM-iso13250-2nd-ed-v2.pdf

Janes, Joseph W. 1994. "Other Peoples´ Judgments: A Comparison of Users and Others´ Judgments of Document Relevance, Topicality, and Utility." *Journal of the American Society for Information Science and Technology* 45: 160-71.

Jarneving, Bo. 2005. "A Comparison of Two Bibliometric Methods for Mapping of the Research Front." *Scientometrics* 65: 245-63.

Krause, Michael G. 1988. "Intellectual Problems of Indexing Picture Collections." *Audiovisual Librarian* 14: 73-81.

Lancaster, Frederick Wilfrid. 2003. *Indexing and Abstracting in Theory and Practice*. Third edition. London: Facet Publishing.

Library of Congress. 2008. *The Subject Headings Manual*. Washington, D.C: Library of Congress, Policy and Standards Division.

Metcalfe, John. 1973. "When is a Subject Not a Subject?" In *Towards a Theory of Librarianship*, ed. Conrad H. Rawski. New York: Scarecrow Press, 303-38.

Miksa, Francis. 1983a. "Melvin Dewey and the Corporate Ideal." In *Melvil Dewey: The man and the Classification*, ed. G. Stevenson and J. Kramer-Greene. Albany, NY: Forest Press, 49-100.

Miksa, Francis. 1983b. *The Subject in the Dictionary Catalog from Cutter to the Present*. Chicago: American Library Association.

Olson, Hope A. 2002. *The Power to Name: Locating the Limits of Subject Representation in Libraries*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Panofsky, Erwin. 1939. *Studies in Iconology: Humanistic Themes in the Art of the Renaissance*. New York: Oxford University Press.

Ranganathan, Shiyali Ramamrita. 1963. *Documentation and its Facets*. London: Asian Publishing House.

Ranganathan, Shiyali Ramamrita. 1964. "Subject Heading and Facet Analysis." *Journal of Documentation* 20: 109–19.

Ranganathan, Shiyali Ramamrita. 1967. *Prolegomena to Library Classification*, 3rd ed. London: Asia Publishing House.

Rowley, Jennifer and Richard Hartley. 2008. *Organizing Knowledge. An Introduction to Managing Access to Information*, 4th ed. Aldershot: Ashgate Publishing Limited.

Salsburg, David. 2001. *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twentieth Century*. New York: W. H. Freeman.

64

Knowl. Org. 44(2017)No.1
Reviews of Concepts in Knowledge Organization

Saracevic, Tefko. 2008. "Effects of Inconsistent Relevance Judgments on Information Retrieval Test Results: A Historical Perspective." *Library Trends* 56, no. 4: 763-83.

Schneider, Jesper W. and Pia Borlund. 2004. "Introduction to Bibliometrics for Construction and Maintenance of Thesauri: Methodical considerations." *Journal of Documentation* 60: 524-49.

Shatford, Sara. 1986. "Analyzing the Subject of a Picture: A Theoretical Approach." *Cataloging & Classification Quarterly* 6, no. 3: 39-62.

Siewert, Charles. 2016. "Consciousness and Intentionality." In *The Stanford Encyclopedia of Philosophy*, Fall 2016 ed., ed. Edward N. Zalta. http://plato.stanford.edu/archives/fall2016/entries/consciousness-intentionality/

Small, Henry G. 1978. "Cited Documents as Concept Symbols." *Social Studies of Science* 8: 327-40.

Soergel, Dagobert. 1974. *Indexing Languages and Thesauri: Construction and Maintenance*. Los Angeles, Calif: Melville Publishing.

Soergel, Dagobert. 1985. *Organizing Information: Principles of Data Base and Retrieval Systems*. Orlando, FL: Academic Press.

Tredinnick, Luke. 2006. *Digital Information Contexts: Theoretical Approaches to Understanding Digital Information*. Oxford: Chandos.

Welty, Christopher A. 1998. "The Ontological Nature of Subject Taxonomies." In *Proceedings of the First Conference on Formal Ontology and Information Systems*, ed. Nicola Guarino. Amsterdam: IOS Press. http://www.cs.vassar.edu/faculty/welty/papers/fois-98/fois-98-1.html

Wilson, Patrick. 1968. *Two Kinds of Power: An Essay on Bibliographical Control*. Berkeley: University of California Press.

Xu, Yuniie and Hainan Yin. 2008. "Novelty and Topicality in Interactive Information Retrieval." *Journal of the American Society for Information Science and Technology* 59: 201–15.

Yablo, Stephen. 2014. *Aboutness*. Princeton, NJ: Oxford: Princeton University Press.