Back to the fundamentals again

A redefinition of information and associated LIS concepts following a deductive approach

Liangzhi Yu
Department of Information Resource Management, The Business School, Nankai University, Tianjin, China

Received 12 December 2014
Revised 12 December 2014
Accepted 14 December 2014

Abstract

Purpose – The purpose of this paper is to redefine information and other associated library and information science (LIS) concepts and to reformulate the mission of the library and information profession and the problem of LIS using these concepts.

Design/methodology/approach – This study adopts a deductive approach to conceptualization, starting from one given, a priori concept.

Findings – This paper develops a constellation of concepts which offer mutual clarification for each other. Having defined data by drawing on its existing denotations, it defines "information" as the combined product of data and meaning, and "document" as the combined product of information and media; it defines "knowledge" as one type of meaning and "work" as one type of information. It shows that the mission of the library and information profession is to ensure maximum discoverability and accessibility of information, and that LIS is structured into two fields correspondingly, each consisting of three tiers of knowledge: philosophical foundations, theories, and technologies.

Practical implications – The redefinition of basic LIS concepts may have practical implications for LIS curriculum design and for the cultivation of professional identity among LIS students in the all-encompassing I-Schools.

Originality/value – This study has formed a coherent conceptual framework for LIS and has clarified the hitherto rather confusing relationship between data, information, and knowledge, and the rather nebulous structure of LIS problems; it sheds some light on the source of conflicts between the subjective and objective conceptualization of information and questions the prevailing understanding of work as ideas or meanings.

Keywords Knowledge, Data, Information, Documents, Works

Paper type Conceptual paper

Introduction

Information has been widely, if not unanimously, regarded as one of the central concepts for library and information science (LIS), not least because the field itself is named after it and presumably takes it as its object of study. However, until now, information has not earned an agreed definition within the field. The first attempt to conceptualize the term specifically for information science dates back to the 1960s (Wellisch, 1972); the latest are still being reported (Budd, 2011; Gnoli and Ridi, 2014; Ma, 2012; Robinson and Bawden, 2014). In between these decades, many different definitions emerged. These definitions are not without impact, but none seem to dominate LIS understanding of information.

A number of reasons may be observed or conjectured for this persistent disagreement. The first and most obvious lies in the fact that information is a term widely used within and beyond academic disciplines (Derr, 1985; Robinson and Bawden, 2014). It is not uncommon to see influences from varied sources manifesting themselves in the definitions of information in LIS. The attempt to define information as uncertainty or the reduction of

The author would like to thank the two referees of this paper for pointing to overlooked literature and for their insightful comments which helped to improve the manuscript enormously.

Journal of Documentation
Vol. 71 No. 4, 2015
pp. 795-816

0022-0418
DOI 10.1108/JD-12-2014-0171

uncertainty (e.g. Wersig, cited in Ingwersen and Jarvelin, 2005), for instance, demonstrates clearly the influence of Shannon and Weaver's information theory; the attempt to define information as the pattern or complexity of organization (e.g. Bates, 2006; Bawden, 2007; Madden, 2004) reveals the influence from theories of evolution; and the attempt to define information as a communicated message or the content of a message (e.g. Shenton, 2004; Svenonius, 2000) shows arguably the influence of the term's usage in ordinary discourse. When different sources of influence fail to find common ground for dialogue, disagreements inevitably arise. The second reason lies probably in the fact that the concept of information is closely associated with such epistemological concepts as knowledge, knowing, and truth, and like these latter concepts, is susceptible to the influence of fundamental philosophical disagreements. This kind of influence is most clearly revealed in the debate between Bates (2008) and Hjørland (2007, 2009) regarding the subjective/objective nature of information, and in the critiques of the DIKW (data, information, knowledge, and wisdom) hierarchy by Frické (2009), Ma (2012), Nielsen and Hjørland (2014), and like-minded scholars. Besides these two reasons, it can be argued that the lack of a solid methodology in defining information has also contributed to the divergences of its denotation. For an elusive phenomenon such as information, it is difficult if not impossible to base its conceptualization on empirical observation, for to define it in this way, we would have to know what to observe first. Confronted with the inaptness of empirical procedures, definitional efforts often resort to personal views rather than alternative methodologies. The literature review section will provide a more detailed reflection on this matter; suffice to note here that, mediated by researchers' epistemological and theoretical perspective and past experiences, personal views are more likely to lead to disagreement than to consensus.

It is the belief of this paper that, by comparison, the methodological hurdle to shared understanding of information is the easiest to tackle and is, therefore, the most promising to start with. With the academic world's deep-rooted tradition to respect methodological rigor, an explicit methodology may offer common ground for conceptual reflection. As a result, shared understanding may gradually develop, probably on the pragmatic if not epistemological level. Driven by this belief, this paper attempts to define information and associated concepts by following an explicitly stated deductive approach. More specifically, this paper starts from one given, a priori concept and arrives at the definitions of other concepts by deducting from this given one. The next (i.e. the second) section reviews the existing conceptualization of information and presents a critique of it from the methodological perspective; the third section explains the deductive approach and the selection of its logical starting point; the sections that follow define, respectively, data, information, and other basic concepts of LIS, each on the basis of the one(s) preceding it. The conclusion section brings together these concepts and discusses the theoretical and practical implications of their redefinitions.

A review and methodological reflection of previous definitions of information

Although predecessors of LIS (librarianship, bibliography, documentation, and information science) had been using the term "information" long before it became a central concept of any of these fields, attempt to formally define it did not emerge until 1965, according to Wellisch (1972). In that year, a group of participants at the symposium on education for information science defined information as "recorded marks" (Wellisch, 1972, p. 171). In 1968, Mikhailov and his colleagues defined information as "the objective content of the link between interacting material objects, which reveals itself in the transformed status of

these objects" (cited in Wellisch, 1972, p. 172). That year also coincided with the transformation of the American Documentation Institute to "the American Society for Information Science."

Attempts to define the term for the newly named information science surged and persisted since 1969, the latest including Zins's (2007) framework of definition based on a Delphi study, Bawden's (2007) definition from the perspective of evolution, and Budd's (2011) definition in terms of meaning and truth. There have also been a large number of review articles that either summarize periodically various definitions advanced, or critique particularly a certain definitional model (Bawden, 2001; Belkin, 1978; Capurro and Hjørland, 2003; Cornelius, 2002; Frické, 2009; Ma, 2012; Meadow and Yuan, 1997; Thellefsen *et al.*, 2014; Wellisch, 1972; Zhang, 1988). This section will not duplicate or merely update these reviews. It focusses, instead, on a comparative review and a methodological critique of the divergent conceptualizations within LIS, thereby demonstrating the degree of conceptual inconsistency facing LIS researchers, educators, and students. For this purpose, it reviews only those definitions proposed within or for LIS.

Table I shows some of the existing conceptualizations of information. Column 1 groups these definitions according to their primary referents. Column 2 compares the essential features of information that these definitions highlight. Column 3 summarizes the major dimensions on which existing conceptualizations diverge, which are also illustrated in Figure 1.

As Table I shows, there exist a wide range of divergences, both fundamental and subtle, in the current conceptualization of information for LIS. The proposed definitions differ first in what they refer to as "information," quoting as varied as all the following as the referents of "information": anything that has the potential to inform or actually informs, things that are formed with the intention to inform, signs used for communication, contents that are being communicated, things that reduce uncertainty, effect of uncertainty reduction itself, and a number of these entities simultaneously (multiple referents). Existing definitions also differ in their assertions regarding the ontological nature (subjective or objective) of information. While some definitions see an entity as information only in so far as a person feels informed by it or regards it as informative, others see it as information as long as it possesses certain objective properties (e.g. the intention to inform).

Existing definitions further differ in the semantic relationships they establish between information and the broader categories (data, knowledge, or message) in terms of which it is defined. Of those who define information in terms of data (rows 2 and 3 in Table I), some see it as a special kind of data (e.g. "data of value in decision making"), some see it as transformed data (e.g. "data processed and assembled into a meaningful form," "the result of processing of data). Of those who define information in terms of message (rows 4 and 5 in Table I), some see it as a type of message; some see it as the content of a message. Similarly, of those who define information in terms of knowledge, some (e.g. Brookes, 1980, p. 131) state that information and knowledge are of the same kind; some (e.g. Nitecki, 1985, p. 387) contend that information and knowledge are different stages of the same continuous process.

As already discussed in "Introduction," there may be various reasons why consensus on defining information is so difficult to obtain. This paper, however, chooses to reflect particularly on the methodological rigor on which existing definitions are based. In general, related studies seem to have abandoned empirical observation as the valid method for defining information, but very few explicitly explain what they have adopted in its place. It is as if existing studies have given up methodological consideration

Table I.
Different conceptualizations of information

Conceptualization clusters	Referents and features	Summary of divergence
1. Define information as the physical representation of what is meant to be communicated, e.g., physical surrogate of knowledge (Farradane, 1979, p. 13), objectivized form of knowledge (reviewed in Bawden, 2001, p. 95); meaningful representation of determinations made of objects (Derr, 1985, p. 496)	Physical representation of knowledge or meaning	Types of referent Certain Uncertain: anything that the receiver regarded as informative Multiple
2. Define information as a qualified type of data, e.g., data of value in decision making (Yovits, 1969, p. 369), data that has been processed and assembled into a meaningful form (Meadows *et al.*, 1982, p. 91); highly concentrated and improved data derived from raw facts (reviewed in Bawden, 2001, p. 95); assemblage of data in a comprehensible form capable of communication and use (International Encyclopedia of Information and Library Science, 1997, p. 184); well-formed meaningful data (Floridi, 2010, p. 21)	A qualified type of data, representation of meaning	Broader category in which information is defined Data Knowledge Message Process Structure A type of its own
3. Define information as the result of processing of data, usually formalized processing (Hayes, cited in Wellisch, 1972, p. 172)	Product of data processing	
4. Define information as a qualified type of message (although there is no consensus on what is meant by message), e.g., messages that convey meaning (Shenton, 2004, p. 370), messages that meet any of the following criteria: (i) truth; (ii) utility; (iii) novelty; (iv) unexpectedness; (v) uncertainty reduction (reviewed in Furner, 2004, p. 440)	A qualified type of message	Realm in which the referent is located Material Non-material realm
5. Define information as the content of a message (Svenonius, 2000, p. 7); that which justifies representational activity (MacKay, cited in Wellisch, 1972, p. 174); the objective content of the link between interacting material objects, which reveals itself in the transformed status of these objects (Mikhailov *et al.*, cited in Wellisch, 1972, p. 172)	Content of message or communication	Essential nature of information Objective, context-independent Subjective, context-dependent
6. Define information as a structure of concepts (Brookes, 1980, p. 131); structure of texts (Belkin and Robertson, 1976, p. 201; Ingwersen and Jarvelin, 2005, p. 10); organization of our experience (Thompson, 1968, p. 305); that which changes or which is capable of changing the image structure or knowledge structure of the recipient (Paisley, reviewed in Meadow and Yuan, 1997, p. 709)	Structure that changes or is capable of changing the recipient's knowledge	

(*continued*)

Table I.

Conceptualization clusters	Referents and features	Summary of divergence
7. Define information as what being recognized as informative in a particular context (Buckland, 1991; Hjørland, 2007, p. 1449; Ma, 2010, p. 4; Budd, 2011, p. 69-70; Cornelius, 2004, p. 385-386); what can answer important questions related to the activities of the target group (Capurro and Hjorland, 2003, p. 390); that reduces an individual's uncertainties relevant to his or her changing purposeful state (Nauta, 1972, cited and agreed by Artandi, 1973, p. 244); system-specific interpretation of external stimuli (Otten cited in Belkin, 1978, p. 69)	A social construct, anything regarded as informative in particular context, anything that produces system-specific effect	
8. Define information as the reduction of uncertainty or doubt and the effect of a message on a recipient (Wersig, cited in Ingwersen and Jarvelin, 2005, p. 12); the amount of complexity to be reduced or that has been reduced (Wersig, 1997, p. 225, cited in Cornelius, 2002, p. 400)	Effect of reduction in uncertainty	
9. Define information as perceptions brought to our attention, but not yet fully assimilated (Nitecki, 1985, p. 388)	Sense data before being assimilated by human mind	
10. Define information as action or event or process during communication, e.g., meaningful communicative action that aims at truth claims and conditions (Budd, 2011, p. 70); an event which takes place at a particular point in time and space to some particular individual that make him or her "in-formed" (Pratt, 1977, p. 215)	Actions or events or series of actions that inform	
11. Define information as a universal property of matter, energy and consciousness, e.g., a stimulus which expands or amends the World View of the informed (Madden, 2004, p. 9); pattern of organization of matter and energy (Bates, 2006, p. 1033); basic property of the universe, that which is contained in the organization of any system (Stonier, 1991); organized complexity in the physical domain, meaning in context in the biological domain, and understanding in the human domain (Bawden, 2007, p. 307); a fundamental property of matter and of consciousness, acting to connect the two by means of its relationship with variety and reflection (Ursul, reviewed in Belkin, 1978, p. 70)	Property of matter, energy, and consciousness, as universal as matter and energy	
12. Define information as a concept with multiple denotations, e.g., Buckland's (1991) triple denotations of information-as-thing, information-as-process, information-as-knowledge; Koblitz's (cited in Wellisch, 1972, p. 173) triple denotations of semantic information as message, semantic information as process, and documentary information	Multiple referents, different features	

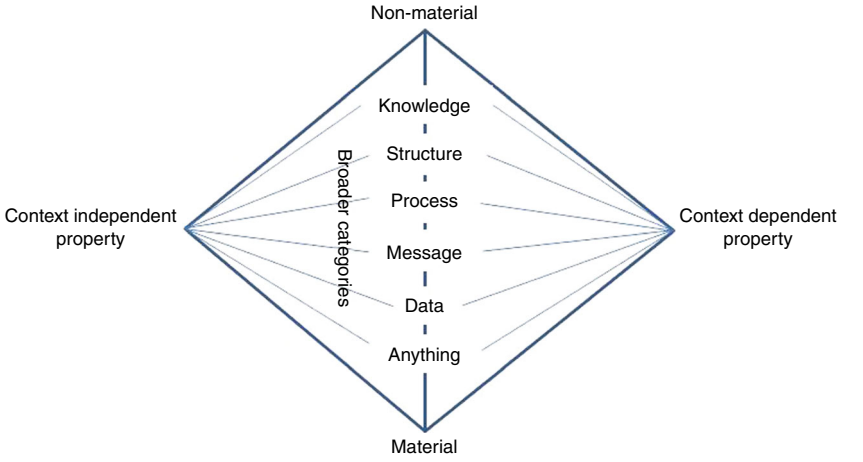

Figure 1.
Divergences of existing conceptualization of information

altogether. It is true that most definitions are supported by various degrees of reasoning. However, it is precisely here that the relaxing of methodological rigor becomes particularly notable.

First, very few studies explicitly state their starting point for reasoning. Where a starting point is indeed given, it often appears disputable. Of those specified starting points, three types can be identified. The first is the linguistic feature (lexical, semantic, etymological, etc.) of the word "information." Pratt (1977, p. 215, also reviewed in Cole, 1994), for example, argued that being composed of "in" and "formation," "information" involves intrinsically changes from within a person. He therefore defined information as an event which takes place at a particular point in time and space to some particular individual that makes him or her "in-formed." This kind of starting point is disputable because whichever linguistic feature is chosen, there are likely to be more than one starting point. Lexically, for instance, the word information is associated with a number of other words like inform, informative, informing, and informed; it is difficult to tell which one is more legitimate than others to serve as the logical starting point for reasoning. The second starting point specified in existing literature is the evolutionary beginning point of the universe, the property of matter and energy. Stonier (1991), Madden (2004), Bates (2006), and Bawden (2007), for example, developed their definitions by showing that the same information "gene" (property or pattern of organization) was within matter and energy right from the beginning of the universe and naturally "inherited" by organisms, animals, and, finally, human beings. This, as these authors contend, makes it possible to establish information as a universal phenomenon and to form a universal definition of information for all disciplines. This starting point is disputable because it leads to a definition of information which is far broader than what LIS has been actually concerned with. The third starting point in existing literature is the pragmatic requirement of information science. Belkin (1978) and Thellefsen *et al.* (2014), for instance, respectively, proposed a set of criteria to guide the conceptualization of information. This starting point is disputable because requirements are essentially concerned with the property of the end product and may not serve well as a starting point for reasoning – besides the fact that different conceptualizations of information science may emphasize very different requirements for the information concept. The absence of an explicit, indisputable, a priori starting point has not only allowed each definitional

effort to go its own way, but also deprived the field of common ground to judge the validity of the resulting definitions. The occasional outbreak of heated debates (Bates, 2008; Hjørland, 2007, 2009) attests to the need of such common ground.

Second, the relaxing of methodological rigor is also reflected in the occasional neglect of established conventions of conceptualization. The context-dependent and the multiple-denotation definitions are particularly notable in this regard. The context-dependent definitions differ from conceptualization conventions because in general, when we define something, we tend to define it by its most essential and stable property and avoid its variable characteristics; otherwise, we risk falling into categorization or nomenclature disorders. In medical science, for instance, disorder would arise if an illness is regarded as a disease for one person and non-disease for another. In the case of LIS, the variability of "information status" also poses great difficulties for theorizing where some degree of comparability, hence objectivity of information is assumed, e.g. when dealing with issues of aboutness, relevance, and information inequality. The multiple-denotation definitions differ from conventions because, in general, the formal definition of a term is a process to eliminate its diverse meanings afforded by its casual use, not a process to reinforce it; otherwise, we risk of confusing ourselves by confounding referents. In the case of information, for instance, upon accepting its multiple denotations, we may began to wonder which referent the term is referring to in such issues as information access, information use, information overload, information poverty, etc.

In summary, because of the close-to-metaphysical status of information, there are few empirical means for either developing or verifying/falsifying definitions for it. The impracticality of the empirical approach seems to have led LIS scholars to relax, perhaps unwittingly, their criteria for methodological rigor and to allow personal views to enter their arguments, which in turn, have led to sustained disagreement. A viable approach to increase shared understanding of the concept, therefore, is to enhance LIS sensitivity to requirements of alternative methodologies, particularly those of the systematic deductive reasoning.

A brief note on the conceptualization approach of this study

In light of the methodological critique of existing studies, this paper adopts a deductive approach in defining information and associated concepts, starting from one given, a priori concept. "Data" is chosen in this paper as the given concept from which the denotations of all other concepts are deduced. This concept is regarded as a qualified starting point for two reasons. First, in comparison with other basic LIS concepts, "data" has arguably enjoyed relatively greater agreement with regard to its denotation. It is generally agreed that data is something to which meaning can be attached. It is hoped that from this very small agreement, further agreement may develop. Second, when data, information, and knowledge are considered in relation to each other, data tend to be considered as antecedent to information and knowledge. So, if a conceptual chain needs to be developed, it has to start from the concept of data.

Having taken the concept of data as given, this paper defines information on the basis of data and defines knowledge on the basis of information. It goes on to define other basic LIS concepts – work, document, the library and information profession, and the LIS discipline – on the basis of data, information, and knowledge together. It is hoped that, on the condition that the concept of data is defined clearly and that the logic of deduction is followed properly, the study will arrive at a constellation of concepts that clarify at least some of the confusion in the existing conceptualization of information and associated concepts.

The concept of data

In comparison with information and knowledge, data has not inspired as much dedicated definitional effort. Most definitions of data are proposed as forerunners to in-depth discussion of information. As already mentioned, LIS scholars tend to agree that: data exist external to a human mind (Zins, 2007); data normally take the format of a set of symbols (Meadow and Yuan, 1997); and as such, data can be attached with meaning, used for communication, and can be manipulated, operated, and processed.

Drawing on these commonly held connotations, this paper defines data as a set of words, numbers, symbols, sound, pictures, and/or other codes of communication that are formed through a composition process. Pictures taken by satellites, video clips taken by a city's street camera, human-computer interactions recorded by screen capture software, words written by a scientist, lines written by a poet, spoken words delivered by a lecturer in class, are all data. They are the product composed of linguistic, numerical, graphical, and/or other signs and are therefore often composed by a human or an artificial intelligent creator through appropriating suitable signs. For instance, within the English language system, there is a word "the," a word "three," and a word "idiot." An Indian scripter Abhijat Joshi chose these three words and composed the phrase "The Three Idiots." In this case, Abhijat Joshi is the producer of this particular phrase, a particular datum. In a similar way, a person or a computer program can produce a table using a set of Arabic numbers, like those in Figure 2.

Defined in this way, data in fact refer to the physical entities either waiting to be assigned with meaning (in the case of satellite pictures) or being considered without reference to the meaning already assigned (in the case of lines written by a poet). Although data can be formed preceding to its being assigned with meaning – modern information technologies are producing this type of data in such a scale that they are now called the big data, yet until now, the primary type of data are still those which are composed with meaning. To put it another way, the intention to convey a specific meaning is still the major reason why data are created by their creators. Taking "The Three Idiots" as an example, Abhijat Joshi chose these three words to name a film. Naming a film is a meaningful action and this action simultaneously gives rise to a datum.

Having noted that the majority of data are born with meaning, it must be emphasized at the same time that data and meaning are separate entities. This is first reflected in the fact that the same data (the same composition of words, numbers, or other symbols) can be created by different persons to convey different meanings. For example, two hostile and competitive teams of three persons may call each other "The Three Idiots," thereby assigning to the three words completely different meanings from what Abhijat Joshi has given them. That data and meaning are separate entities is also reflected in the fact that data can be processed and manipulated independently without reference to the meaning they convey. A search engine, for example, can search the database of web pages for all pages where the

90	97	95	97
91	82	75	88

Figure 2.
A hypothetical set of data

phrase "The Three Idiots" is stored as data. It does so by matching the words without considering their meaning. It therefore may return web pages with the required data but not the required meaning.

This concept of data is similar to but not identical with Furner's (2004) concept of utterance and Ingwersen and Jarvelin's (2005, p. 8) concept of intentional signs. Furner defined an utterance (or expression) as "a particular vehicle of meaning" and a "product of a human decision to act by using words in a way that has meaning" (Furner, 2004, p. 431). Ingwersen and Jarvelin also see intentional signs as products of directed human action in particular context. However, as the aforementioned examples of data have already illustrated, such utterances or signs only make a part, albeit a large part, of all data that exist. An increasing amount of data is now being produced by information technologies without a particular intention.

Definition of information based on data

As already noted, within and without LIS, data has been widely regarded as antecedent to information. Having defined data as a set of words, numbers, symbols, sound, pictures, and/or other signs that are formed through a composition process, it is opportune to ask how data can become informative or how data can be transformed into information. Intuitively we know that for something to be informative, it has to be meaningful. Therefore, for data to be informative, it has to be married with meaning. It is possible now to define information as the combined product of data and meaning. So a certain piece of information is a certain set of words, numbers, symbols, sound, pictures, and/or other signs combined with a certain meaning. Therefore, while the numbers in Figure 2 make a set of data, the same set of numbers in Figure 3 together with the meaning attached to it makes a piece of information. By the same token, the set of words "The Three Idiots" expressing the name of an Indian film makes one piece of information while the same set of words expressing one team's hostility toward its rival team makes another. Thus defined, information can be expressed as: "Information = Data+Meaning."

It is already known that data are composed of linguistic, numerical, graphical, and other types of communication materials (words, numbers, symbols, pictures, sounds, signals, etc.). So the above formula of information can be further developed into: "Information = {words, numbers, symbols, pictures, sounds or signals, etc.} +Meaning."

As Furner (2004) noted, "meaning" is also a complex concept in LIS that requires definition. For the purpose of conceptualizing information, however, we can perhaps bypass the task of defining meaning and rely on our experience to specify a few of its illustrative types. When people deploy certain communication materials to express something, what they express can be facts, news, knowledge, ideas, imaginations, conjectures, etc. Taking these as illustrative categories of meaning, the formula of

Categories of students	Math	Reading	Writing	Oral
Students with reading habit	90	97	95	97
Students without reading habit	91	82	75	88

Comparison of test scores between students with or without reading habit

Figure 3.
Hypothetical information formed by assigning meaning of the data in Figure 2

information can be further expressed as: "Information = {words, numbers, symbols, pictures, sounds or signals, etc.}+{facts, news, knowledge, ideas, imaginations, conjectures, etc.}."

The above formula brings about a new question crucial to the understanding of information: How are words, numbers symbols, etc., combined with facts, news, knowledge, or other types of meaning to form information? For any given event, there may be different versions of news made possible by employing different sets of words, numbers, symbols, etc.; similarly, for any given phenomenon, there may be different kinds of explanations, again made possible by different sets of words, numbers, symbols, and/or other types of data. To those Chinese people who participated in the 1999 protest against the US bombardment of the Chinese embassy in the former Yugoslavia, for instance, different media applied drastically different labels, e.g., patriotic protesters, anti-US protesters, and mobs, giving rise to very different meanings for the same protest. The resulting information attests to the great complexity of unification between data and meaning in relation to a rather hazy external reality, the event itself. As a large number of LIS scholars (e.g. Day, 1998, 2000, 2008; Frohmann, 1992; Talja *et al.*, 2005) repeatedly argue, the combination between data and meaning does not simply happen in a neutral, objective, and mechanical manner, but is mediated very much by the social, cultural, political, institutional, professional, and other contexts in which the combination takes place. Contemporary philosophers such as Foucault contend that even knowledge takes its form within certain discursive formation (Radford, 2003). To reflect this "constructedness" of information, the aforementioned formula can be further developed into: "Information = {words, numbers, symbols, pictures, sounds or signals, etc.}+{facts, news, knowledge, ideas, imaginations, conjectures, etc.}, where both {facts, news, knowledge, ideas, imaginations, conjectures, etc.} and their expression in {words, numbers, symbols, pictures, sounds or signals, etc.} are formed in specific contexts from specific standpoints in relation to an external reality."

Defined in this way, information is bound to have complex attributes associated with both its data and meaning. It has, for instance, a language attribute associated with its data and an "aboutness" attribute associated with its meaning. Like attributes of any other things, attributes of information have values, in the same way as the attribute of sex has either male or female as its value. Table II lists some illustrative attributes of information and their possible values. These are based on cursory observations and do not claim to exhaust all information attributes; enumeration of a complete list of such attributes is beyond the scope of this paper.

Much of the library and information work deals with attributes of information. It is through the process of assigning appropriate values to the attributes of certain information that library and information professionals produce representations or surrogates of the information, which when aggregated with each other, form information retrieval systems. This is why attributes of information and the way to effectively assign values to them have formed a core area of study in LIS. Cataloguing rules, classification schemes, subject heading lists, and thesaurus are all inventions of LIS within this area. Despite the continued advancement of these technologies, the library and information profession is yet to find effective ways to deal with some of the most complex attributes of information. One of such attributes is the "aboutness" of information. It has long been recognized that different people in different contexts can assign very different values to the "aboutness" attribute of any given information and, for this reason, there is no guarantee that the author, information professionals, and different users of certain information will assign the same value to the same

Attribute name	Meaning of the attribute	Possible values
Format	The systems of signs that the data of information takes	Words, numbers, symbols, pictures, sounds or signals, etc.
Language	The language used when the value of the format attribute is "words"	Chinese, English, French, Italian, Russian, etc.
Quantity	Size of the data	Number of words or bits or other similar measurement
Types of meaning	Categories of meaning conveyed by the data	News, fact, knowledge, imagination, conjectures, etc.
Discipline	Academic areas for which the meaning is meant when type of meaning is knowledge	Class from a classification scheme
Aboutness	Topics, events, persons, etc., to which the meaning is dedicated	Subjects from a subject heading list or a concept from an ontology
Informativeness	Degree of details that the information tells about the topics, events, persons, etc.	Measurement yet to be developed
Perspectives	The standing point from which the meaning becomes meaningful	Political or theoretical views
Reliability	Truthfulness of meaning when its type is fact, news, and knowledge	True or false
Novelty	Originality or newness in comparison with other information on the same "aboutness"	"Reviews, textbooks, discoveries" for knowledge
Creator	The person or corporate body who combined the data and meaning	Name of creators
Time of creation	Time when data and meaning is combined to form information	Date

Table II.
Illustrative attributes of information

information. An intriguing question facing library and information professionals then is whether they should be as loyal as possible to the author or to the user, and how? Another complex attribute is the "informativeness" of information which means the amount of news, facts, ideas, knowledge, etc., that any given information provides regarding any given event, phenomenon, relationship, etc. Intuitively we know that the following pieces of information about the same event differ in degree of informativeness: first, "there was an accident on M1 this morning"; and second, "at 6 a.m. this morning, a lorry crashed into a car on M1 which killed two people." We also know that the value of this attribute, if properly assigned, would be a great help to those who seek information. However, important as it is, the way to assign value to this attribute is still unknown to LIS. Shannon's measurement of information quantity offers little help because it does not deal with meaning.

It can be argued that the more library and information professionals know about and make use of the attributes of information, the greater is their capacity to organize information for effective search. No information search tools can claim to be fully functional unless they have fully explored the utility of attributes of both data and meaning. It is fitting to note here that despite the great popularity of Google and similar search engines, they have so far made little use of the meaning attributes of information and have been apparently handicapped by this indifference to meaning: they can only match web pages with search queries based on data attributes, which almost always return very large proportion of irrelevant items. This is why the semantic web and linked data technologies, which take both data and meaning attributes of information into consideration, are seen to hold far greater promise than the current search engines in improving information search.

Having defined information and its attributes in this way, it is now possible to compare this definition with previous conceptualizations of information. Three differences are worth special note. First, while many previous conceptualizations define anything that is informative to be information, this study, however, contends that things, events, phenomena, etc., are not information in themselves, even though they can become informative in particular contexts; descriptions of these things, events, and phenomena using words, numbers, symbols, etc., however, are. To put it another way, as far as LIS is concerned, things, events, and phenomena are only sources of information, not information itself. Second, while some of the previous definitions see the status of something being information as context-dependent, the definition of this study, however, contends that information acquires its status whenever certain meaning is combined with certain data and that the information status of the combined product does not change with context. Third, while some of the previous definitions define information as a species or a qualified type of data, in forms such as "Information is data of X type," the definition of this study emphasizes that information and data are entities of a different kind; data being only one ingredient of information.

Definition of knowledge, works, and documents based on data and information

In defining the attributes of information, this paper has already categorized knowledge as one type of meaning of information. It is time to further explicate and justify this categorization. We can do this by examining indisputable examples of knowledge. Even before a formal definition of knowledge is proposed, we would perhaps readily accept that Newton's law of universal gravitation is a part of human knowledge and that it remains the same knowledge regardless of the format (formula, graph, or words) or language in which it is expressed. In a similar vein, we would also readily accept that Marx's theory of surplus value remains the same knowledge whatever languages have been used to present it. In both cases, the knowledge is what is expressed by formulas, graphs, words, etc., and in both cases, it has to be expressed in this way to be disseminated. As it is already noted, what is expressed by formula, graphs, words, sounds, and symbols is in fact meaning. Knowledge is therefore a type of meaning, parallel with news, imaginations (stories), etc. Being a type of meaning, knowledge is constitutive of information, together with the data that convey it. Newton's law of universal gravitation can be expressed technically (e.g. in formula) in textbook for physics students, or literally in encyclopedia for the general public or figuratively in picture books for children. The same knowledge expressed in different formats for different audiences gives rise to different pieces of information.

The primary difference between knowledge and other types of meaning lies in its production process. In modern society, the most typical means of producing knowledge is through research. There are of course other ways to produce knowledge, but researchers are generally seen as professional knowledge creators by training and by division of labor and are therefore the main source of knowledge in modern society.

Using the terms provided by the concepts of data and information, we can now define knowledge as one type of meaning conveyed by data and constitutive of information, whose production usually involves substantial intellectual labor, where the criteria for being substantial is normally agreed and authorized by the research community.

This brings in another LIS concept associated with intellectual activities, work. In LIS, the concept of work is generally used to refer to the intellectual creation of an author (Svenonius, 2000). As such, works have been commonly regarded as different

from their physical embodiment, documents. But how should a work be aligned with data, meaning, information, and knowledge? Is it of the same kind as any of these or is it a type of its own? Perhaps the best way to relate works with these concepts is to formulate unquestionable statements about them and then examine the relationships as revealed by these statements:

Statement 1. The work "Dream of the Red Chamber" by Cao Xueqin tells a story about the rise and decline of an aristocratic family in the height of feudal China.

Statement 2. It is impossible for Cao Xueqin to think of, i.e., to create "Dream of the Red Chamber" without the Chinese language.

Statement 3. No works exist without being expressed in words, graphs, sound, or other expressive signs.

Statement 4. The work "Dream of the Red Chamber" by Cao Xueqin has produced a great many editions of books.

Statement 1 involves the relationship between a work and its meaning and shows that, as work contains meaning (the story), it cannot possibly be of the same kind as meaning. Statements 2 and 3 involve the relationship between a work and its data, and show that work is also a different kind from data, but depends on data for its creation. Statements 1, 2, and 3 together involve the relationship between a work and its data and meaning at the same time, and show that both data and meaning are constitutive of work. As work consists of data and meaning in the same way as information, it can be said to have the same genetic origin as information. It is therefore safe to say that works are of the same kind as information. Moreover, being a combination of a relatively complete set of meaning and a definable set of data, a work makes an easily distinguishable unit of information. This is why the library and information profession tend to take a work as the basic unit of representation in information organization.

Within LIS, work is generally differentiated from other types of information primarily by its nature as an intellectual product (Svenonius, 2000). As already noted, when a piece of information involves substantial intellectual labor, to such an extent that the amount is approved by the research community, the meaning of that information can be categorized as knowledge. For this reason, work appears to make a primary category of information whose meaning is knowledge. To put it another way, knowledge as a type of meaning is most often associated with that type of information called work. Being such a sub-category of information, work shares the same property and attributes of information. In particular, work embeds within it the same "constructedness" and interpretability as information. It is due to these properties that Day (2008) warns against the mechanistic view of works as objective containers of meaning.

Such a conceptualization of work differs markedly from that adopted by Functional Requirements for Bibliographic Records and Resource Description and Access. These two frameworks have empathetically separated work from expressions (language) and have therefore defined work as the content, ideas, or thoughts. Defined as such, work becomes the same kind as meaning.

Having defined data, information, and work, it becomes relatively easy to define document. If information (and work as its sub-category) is the combined product of data and meaning, then document is the combined product of information and a

material media (e.g. paper, disc). Therefore, we have "Document = Information+Material Media" or "Document = {words, numbers, symbols, pictures, sounds or signals, etc.} +{facts, news, knowledge, ideas, imagination, conjunctures}+media, where both {facts, news, knowledge, ideas, imaginations, conjectures, etc.} and their expression in {words, numbers, symbols, pictures, sounds or signals, etc.} are formed in specific contexts from a specific standpoint in relation to an external reality."

A comparison between the basic concepts of LIS

Having defined the basic concepts of LIS, a comparison between their referents, i.e., the phenomena of data, information, knowledge, and documents, in terms of their most fundamental characteristics, is in order. Table III shows these characteristics on three dimensions: ontological, epistemological, and operational.

As shown in Table III, in terms of their forms of existence in the world, both data and information assume existence as signs, except that information has meaning attached to the signs; in contrast, knowledge exists as ideas and document as materials. As signs themselves without meaning do not tell people anything, data in and of itself is not informative. Information is informative because it contains meaning, and document is informative because it carries information. Knowledge, being the product of intellectual activities, is supposed to provide explanation for physical, social, and our inner worlds. It is therefore not merely informative, but also enlightening. However, because of its ontological characteristics, it cannot be directly processed by machine or human hands; neither can it be transmitted directly from one place to another. For knowledge to be processed and transmitted, it first needs to be expressed in words, formulas, symbols etc., and assumes an existence in texts, graphs, and other data formats. However, when this happens, the resulting product is no longer knowledge itself, but information. Therefore, knowledge can only be organized, processed, and transmitted through the organization, processing, and transmission of information. When we talk about "knowledge organization," we are in fact talking about information organization.

It follows that apart from knowledge, all other phenomena behind the basic LIS concepts – data, information, works, and documents – are possible to become the object of LIS activities. In theory, any data, information, and documents can become such objects; in practice, however, LIS professionals have always focussed on works and work-based documents. Whether this focus will change with the emergence of the big data and semantic web is an issue of increasing dispute (Nielsen and Hjørland, 2014).

Mission of the library and information profession based on information, work, and document

As information contains meaning, it informs. It is for this reason that information is needed in all different contexts of human life, e.g., for decision making, problem solving, education, or understanding things. This need often results in information-seeking

Table III.
Ontological, epistemological, and operational characteristics of data, information, knowledge, and document

	Data	Information	Knowledge	Document
Ontological	Signs, symbols	Signs, symbols	Ideas	Materials
Epistemological	Not informative	Informative	Enlightening	Informative
Operational	Transmittable and retrievable	Transmittable and retrievable	Transmittable and retrievable via information	Transmittable and retrievable

behavior, which, in theory if not in practice, can be further divided into two parts: first, to identify whether relevant information exists and where such information is contained and held; and second, to obtain information for use. The former can be called information search and the latter information access.

The fact that neither of these processes is easy was recognized by our ancestors at a very early stage of human history. Libraries and other document centers were established by ancient civilizations to facilitate these processes as early as thirtieth century BC, but difficulties became more acute each time information-related technologies (e.g. paper-making, printing) made break-through in information production. During the seventeenth century, particularly in scientific research and education, information production reached such a scale that its search and access became increasingly challenging. Complex division of labor – the modern scientific communication system – began to take shape in this context. Within this system are: first, players who are entrusted with the responsibility to ensure that scientific information indeed contains reliable, original, and enlightening knowledge and that it has used appropriate data to present the knowledge (peer reviewers); second, players who engage in mass production of documents to disseminate the information (publishers); third, players who process information to form searching tools (bibliographers, indexers, and librarians); and finally, players who gather and organize information for access (librarians). This whole system can be seen as consisting of two tiers. While the upper tier concentrate on information and document production, the lower tier focussed on ensuring information discoverability and accessibility.

It is the librarians who first claimed a professional responsibility for the lower tier. Sadly their capacity for the mission did not keep pace with either the growth of information or the demand of information users, particularly users of the research community. They had been too engrossed with books to give due attention to works contained in journals, proceedings, and anthologies. At the turn of the twentieth century, unsatisfied with librarians' services for information search and access (particularly search), members of the research community began to do the job themselves. The first few decades of the twentieth century saw an increasing number of researchers diverging from their own research activities to engage in information-related research and practices, focussing particularly on the organization of scientific information. They called themselves documentalists and later information scientists. The older library profession would not deny that throughout the twentieth century, their new rivals made tremendous contributions to improving the effectiveness of information search by studying "information storage and retrieval." However, with their focus on information storage and retrieval, information scientists made less impact on information access, a process depending very much on services of various libraries. The mission of ensuring both information search and information access is made whole by the library profession and the information scientists together. In this sense, the term "library and information profession" is indeed the most appropriate title for the whole division of labor concerned with ensuring information search and access (i.e. the lower tier of the communication system). The connection between data, information, knowledge, document, and the mission of library and information profession is shown in Figure 4.

Seen from this definition of the library and information profession, not only the development of traditional search tools such as library catalogues, bibliographies, indexes, and the associated technologies (e.g. classification schemes and thesaurus), but also that of web search engines, semantic webs, and their associated technologies (e.g. ontology and linked data), fall perfectly well within the boundary of the first mission of

the library and information profession, as they all aim at maximizing information discoverability, hence the efficiency and effectiveness of information search. Moreover, as already discussed in the "Definition of information" section, notwithstanding the great popularity of Google in searching web information, it is far from reaching the limit of what the library and information profession can do in enabling information search, primarily because it fails to make use of the meaning attributes of information. By the same token, not only the development of physical collections and associated services in libraries, but also the development of virtual platforms and portals on the internet, fall within the boundary of the second mission of the library and information profession, because they all aim at ensuring maximum information access. Notwithstanding the rapid diffusion of all kinds of digital technologies for information access, the need to ensure efficient, and effective access remains as acute as ever. This is at least in part because new technologies have created as many obstacles to information access (e.g. skill requirements, steep learning curve, dependence on fast evolving hardware, incompatibility of hardware and software) as they have opened opportunities. It is clear that the mission of the library and information profession corresponds to a very broad and dynamic territory of human needs.

Up until now, libraries have been the central platform for both information search and access, hence a central site for activities of the library and information profession. However, beginning from the second half of the twentieth century, this position has been under constant challenges, first by the blossom of bibliographic database producers and vendors, who dealt with information search outside the library walls, and then by the emergence of the internet which facilitated both information search and access outside libraries. It is still too early to tell how physical libraries as information search and access sites will evolve in the future, but neither challenge should be seen as a threat to the library and information profession. A profession disappears only when its mission becomes irrelevant to society or all activities required by its mission can be automated. As the above discussion reveals, neither of these possibilities seem to be likely as far as the mission of library and information profession is concerned.

The structure of LIS based on the mission of the profession

Ensuring the effectiveness of information search is an extremely challenging task which incessantly raises new research questions. This task requires solid understanding of not only the complexity of both data and meaning for the purpose of representing information, but also the need, the cognition, meaning interpretation, and relevance judgment of users. It also requires the field to monitor closely the development of information technologies and apply them to information organization. Moreover, it requires the field to ground its understanding of both information and users and technologies in the philosophical study

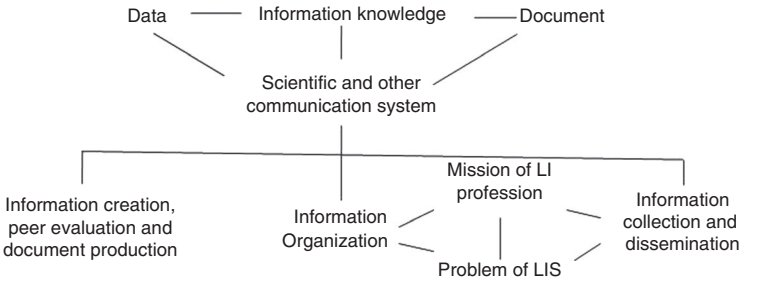

Figure 4.
The relationship between LIS basic concepts and its mission and problem

of information, knowledge, and knowing, i.e., in epistemology. The complexity of issues facing this section of LIS is perhaps best illustrated by the mind-boggling nature of aboutness and relevance of information.

It is no less a challenging task to ensure the effectiveness of information access. This task requires solid understanding of not only the flow of information from its source to its recipients and the play and inter-play of economics, politics, and power behind this flow, but also the cognitive and behavioral characteristics of users in their historical, social, cultural, professional, and organizational contexts. It also requires the field to understand the working of various information access platforms, not least the service and management of libraries. Moreover, it requires the field to ground all its above understanding in epistemological and ethical studies of information. The complexity of issues facing this section of LIS is best illustrated by the many competitive explanations (e.g. cognitive, social-cognitive, and political economy) of why people differ so much in information access.

All these suggest that corresponding to the two missions of the library and information profession, LIS is also structured into two primary fields, dealing, respectively, with issues associated with information search and those associated with information access. Each field is consisted of three tiers of knowledge ranging from philosophical foundations, theories, to technologies/methods. The library and information profession relies on both fields and all three tiers to efficiently and effectively fulfill its mission. The structure is shown in Figure 5. There are of course other areas of inquiry which are not directly related to information search or access, e.g., studies of the profession itself, LIS education, the publishing industry, etc. As these do not produce direct answers to information search and access problems, they form arguably the marginal or inter-disciplinary fields of LIS.

In summary, the discipline of LIS sees information as comprising both data and meaning and therefore being intrinsically informative. It aims to equip the library and information profession with philosophies, theories, and technologies to ensure maximum information search and access, which it believes will in turn maximize the intrinsic value of information (i.e. being informative) for mankind. This leads to a body of knowledge that differs markedly from that advanced by other disciplines, particularly management science and economics. Management science, for instance, sees information as a potentially profitable asset of a particular organization and has developed a body of knowledge to help organizations (especially business organizations) to capitalize on this asset. This body of knowledge includes, but is not limited to, competitive intelligence, information resource

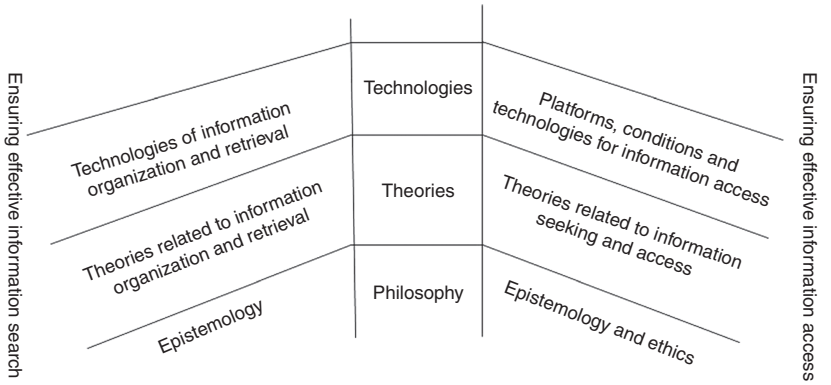

Figure 5.
The structure of the LIS domain

management, and knowledge management. Economics sees information as a production means or commodity and has therefore developed a body of knowledge to explain how information exchange contributes to economic growth. Although their interest in information overlaps that of LIS, neither of these two fields lends pertinent perspectives and knowledge to the mission of library and information profession.

Conclusion

Based on the need for a coherent and unambiguous constellation of concepts for LIS, this paper takes the field's least disputed concept of data as given and defines deductively all other concepts by relating them to data and to each other. In this way, it arrives at a group of definitions for basic LIS concepts that provide mutual clarification for each other. Particularly, it redefined data as a set of signs composed of words, numbers, symbols, pictures, sound, and/or other codes of communication; information as the combination of data and meaning where both the meaning and its expression in data are formed in specific contexts from specific standpoints in relation to an external reality. It then redefined knowledge as a type of meaning of information, works as a species of information, and documents as the combination of information and media. A number of these definitions differ dramatically from previous ones. In particular, the definition of knowledge broke away from previous conceptualizations that see information and knowledge as of the same kind or as transformable from each other; the definition of work broke away from previous conceptualizations that see it as the same kind to meaning or knowledge.

Having been defined in relation to each other, these concepts together provide a conceptual framework to understand a range of fundamental issues in LIS. First, they enable us to compare the characteristics of the phenomena to which they refer. This in turn, establishes information and document as the primary objects of library and information professional activities, and explains why a work has been the primary unit of information. Second, they enable us to understand the mission of the library and information profession and explain why librarians and information scientists do indeed belong to the same profession. Seeing information as the combination of data and meaning, the library and information profession aims to ensure maximum information search and access so as to maximize information's intrinsic value for mankind. The profession fulfills the first part of its mission by processing and organizing information to form various information searching tools; it fulfills the second part by collecting, maintaining, and disseminating information to facilitate information access and by promoting policies and technologies congenial to maximum information access. As neither of these tasks is easy, the library and information profession relies on a correspondingly structured body of knowledge to guide its activities, which forms the domain of LIS. This gives rise to a distinct professional and scientific field which overlaps with but differs essentially from other fields holding the flag of information, e.g., information resource management in management science, information economics in economics, and information technology in computer science. Connecting the problem of LIS to the conceptual framework developed in this study does seem to help clear some of the confusions regarding the identity of LIS recently summarized by Hjørland (2014).

In addition to helping LIS educators to lay a coherent foundation for students to build up their professional knowledge, the definitions advanced by this study may also have a number of other practical implications for LIS education. For one thing, it urges the curriculum design of LIS to cover thoroughly both the two sub-fields and all the three tiers of the body of LIS knowledge to provide students with adequate capacity to

ensure information search and access. For another, it calls on the all-encompassing I-Schools to strengthen the sense of professional identity among students who choose services in information search and access as their own career. Formed in 2004, the I-School movement has incorporated into the same school a number of information-related disciplines and has tried aggressively to forge a new academic territory and identity for itself by pledging to connect information, technology, and people (King, 2006). However, apart from promising a brighter future for LIS, it has so far said little about how it will manage the field's already distinctive mission. Will it reinforce LIS mission to retain the field's distinctive identity within I-School, or will it transform LIS to make it more like other parts of the I-School to the extent that LIS loses its current clarity of mission and identity altogether? If LIS were eventually taken by the latter prospect, would it be too much an exaggeration to compare the I-School movement to LIS's wolf-grandmother?

The conceptualization of LIS concepts in this paper has also raised some further questions for both the researchers and practitioners of LIS. One of these questions regards the definition of the appropriate unit of information in contexts such as information organization and retrieval. Concerns with the unit of information in LIS date back to the early twentieth century and were voiced by highly influential librarians/documentalists such as Kaiser (Dousa, 2014), but the question was rarely taken up in practice, partly because the field has always taken one work as the de facto unit of information. However, is a work the only possible unit of information? Is it the most useful unit for maximum information search and access? When two different sets of data (e.g. a text and a graph) combine with the same meaning, does it make two units of information or one? How will size of unit affect the efficiency and effectiveness of information search and access? In light of the concepts developed in this paper, and in view of the possibility for linked data to connect any part of any work (Gradmann, 2014), these questions have certainly gained new momentum. Another question regards the delineation of information attributes. As this paper has shown, attributes of information concern both data and meaning, each offering some unique utilities for information organization and search. Exactly what attributes does information have? How can appropriate values be assigned to different attributes, particularly meaning-related attributes such as "aboutness" and "informativeness?" Answers to these questions may expose LIS to further possibilities of improving information search and/or access. In any case, a clearly defined constellation of concepts cannot only strengthen the disciplinary status of LIS, but also generate greater usefulness for LIS professional practices.

References

Artandi, S. (1973), "Information concepts and their utility", *Journal of the American Society for Information Science*, Vol. 24 No. 4, pp. 242-245.

Bates, M.J. (2006), "Fundamental forms of information", *Journal of the American Society for Information Science and Technology*, Vol. 57 No. 8, pp. 1033-1045.

Bates, M.J. (2008), "Hjørland's critique of Bates' work on defining information", *Journal of the American Society for Information Science and Technology*, Vol. 59 No. 5, pp. 842-844.

Bawden, D. (2001), "The shifting terminologies of information", *Aslib Proceedings: New Information Perspectives*, Vol. 53 No. 3, pp. 93-98.

Bawden, D. (2007), "Organised complexity, meaning and understanding: an approach to a unified view of information for information science", *Aslib Proceedings: New Information Perspectives*, Vol. 59 Nos 4/5, pp. 307-327.

Belkin, N.J. (1978), "Information concepts for information science", *Journal of Documentation*, Vol. 34 No. 1, pp. 55-85.

Belkin, N.J. and Robertson, S.E. (1976), "Information science and the phenomenon of information", *Journal of the American Society for Information Science*, Vol. 27 No. 4, pp. 197-204.

Brookes, B.C. (1980), "The foundations of information science part I. Philosophical aspects", *Journal of Information Science*, Vol. 2 Nos 3-4, pp. 125-133.

Buckland, M.K. (1991), "Information as thing", *Journal of the American Society for Information Science*, Vol. 42 No. 5, pp. 351-360.

Budd, J.M. (2011), "Meaning, truth, and information: Prolegomena to a theory", *Journal of Documentation*, Vol. 67 No. 1, pp. 56-74.

Capurro, R. and Hjørland, B. (2003), "The concept of information", *Annual Review of Information Science and Technology*, Vol. 37 No. 1, pp. 343-411.

Cole, C. (1994), "Operationalising the notion of information as a subjective construct", *Journal of the American Society for Information Science*, Vol. 45 No. 7, pp. 465-476.

Cornelius, I. (2002), "Theorizing information for information science", *Annual Review of Information Science and Technology*, Vol. 36 No. 1, pp. 392-425.

Cornelius, I. (2004), "Information and its philosophy", *Library Trends*, Vol. 52 No. 3, pp. 377-386.

Day, M.T. (1998), "Transformational discourse: ideologies of organizational change in the academic library and information science literature", *Library Trends*, Vol. 46 No. 4, pp. 635-667.

Day, R.E. (2000), "The 'conduit metaphor' and the nature and politics of information studies", *Journal of the American Society for Information Science*, Vol. 51 No. 9, pp. 805-811.

Day, R.E. (2008), "Works and representation", *Journal of the American Society for Information Science*, Vol. 59 No. 10, pp. 1644-1652.

Derr, R.L. (1985), "The concept of information in ordinary discourse", *Information Processing and Management*, Vol. 21 No. 6, pp. 489-499.

Dousa, T.M. (2014), "Documentary languages and the demarcation of information units in textual information: the case of Julius O. Kaiser's systematic indexing", in Ibekwe-SanJuan, F. and Dousa, T.M. (Eds), *Theories of Information, Communication and Knowledge, Studies in History and Philosophy of Science 34*, Springer-Verlag Berlin Heidelberg, pp. 297-323.

Farradane, J. (1979), "The nature of information", *Journal of Information Science*, Vol. 1 No. 1, pp. 13-17.

Floridi, L. (2010), *Information: A Very Short Introduction*, Oxford University Press, Oxford.

Frické, M. (2009), "The knowledge pyramid: a critique of the DIKW hierarchy", *Journal of Information Science*, Vol. 35 No. 2, pp. 131-142.

Frohmann, B. (1992), "The power of images: a discourse analysis of the cognitive viewpoint", *Journal of Documentation*, Vol. 48 No. 4, pp. 365-386.

Furner, J. (2004), "Information studies without information", *Library Trends*, Vol. 52 No. 3, pp. 427-446.

Gnoli, C. and Ridi, R. (2014), "Unified theory of information, hypertextuality and levels of reality", *Journal of Documentation*, Vol. 70 No. 3, pp. 443-460.

Gradmann, S. (2014), "From containers to content to context: the changing role of libraries in eScience and eScholarship", *Journal of Documentation*, Vol. 70 No. 2, pp. 241-260.

Hjørland, B. (2007), "Information: objective or subjective/situational?", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 10, pp. 1448-1456.

Hjørland, B. (2009), "The controversy over the concept of 'information': a rejoinder to professor Bates", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 3, p. 643.

Hjørland, B. (2014), "Information science and its core concepts: levels of disagreement", in Ibekwe-SanJuan, F. and Dousa, T.M. (Eds), *Theories of Information, Communication and Knowledge, Studies in History and Philosophy of Science 34*, Springer-Verlag, Berlin Heidelberg, pp. 205-235.

Ingwersen, P. and Järvelin, K. (2005), "The sense of information: understanding the cognitive conditional information concept in relation to information acquisition", in Crestani, F. and Ruthven, I. (Eds), *CoLIS 2005, LNCS 3507*, Springer-Verlag, Berlin Heidelberg, pp. 7-19.

International Encyclopedia of Information and Library Science (1997), "Information", in Feather, J. and Sturges, P. (Eds), *International Encyclopedia of Information and Library Science*, 1st ed., Routledge, London, p. 184.

King, J.L. (2006), "Identiy in the I-School movement", *Bulletln of the American Society for Information Science and Technology*, Vol. 32 No. 4, pp. 13-15.

Madden, A.D. (2004), "Evolution and information", *Journal of Documentation*, Vol. 60 No. 1, pp. 9-23.

Ma, L. (2010), "Information as discursive construct", paper presented at the American Society for Information Science and Technology (ASIST) 2010, October 22-27, Pittsburgh, PA, available at: http://onlinelibrary.wiley.com/doi/10.1002/meet.14504701098/pdf (accessed October 10, 2013).

Ma, L. (2012), "Meanings of information: the assumptions and research consequences of three foundational LIS theories", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 4, pp. 716-723.

Meadow, C.T. and Yuan, W. (1997), "Measuring the impact of information: defining the concepts", *Information Processing & Management*, Vol. 33 No. 6, pp. 697-714.

Meadows, A.J., Gordon, M. and Singleton, A. (1982), *Dictionary of New Information Technology*, Kogan Page Ltd, London.

Nielsen, H.J. and Hjørland, B. (2014), "Curating research data: the potential roles of libraries and information professionals", *Journal of Documentation*, Vol. 70 No. 2, pp. 221-240.

Nitecki, J.Z. (1985), "The concept of information-knowledge continuum: implications for librarianship", *Journal of Library History*, Vol. 20 No. 4, pp. 387-407.

Paisley, W. (1980), "Information and work", in Dervin, B. and Voigt, M.J. (Eds), *Progress in Communication Sciences Vol. 2*, Ablex Press, Norwood, NJ, pp. 118-165.

Pratt, A.D. (1977), "The information of the image: a model of the communication process", *Libri*, Vol. 27 No. 3, pp. 204-220.

Radford, G.P. (2003), "Trapped in our own discursive formation: an archaeology of library and information science", *Library Quarterly*, Vol. 73 No. 1, pp. 1-18.

Robinson, L. and Bawden, D. (2014), "Mind the gap: transition between concepts of information in varied domains", in Ibekwe-SanJuan, F. and Dousa, T.M. (Eds), *Theories of Information, Communication and Knowledge, Studies in History and Philosophy of Science 34*, Springer-Verlag, Berlin Heidelberg, pp. 121-141.

Shenton, A.K. (2004), "Operationalising the concept of 'information' for research into information behaviour", *Aslib Proceedings: New Information Perspectives*, Vol. 56 No. 6, pp. 367-372.

Stonier, T. (1991), "Towards a new theory of information", *Journal of Information Science*, Vol. 17 No. 5, pp. 257-263.

Svenonius, E. (2000), *The Intellectual Foundation of Information Organization*, The MIT Press, Cambridge, MA.

Talja, S., Tuominen, K. and Savolainen, R. (2005), "'Isms' in information science: constructivism, collectivism and constructionism", *Journal of Documentation*, Vol. 61 No. 1, pp. 79-101.

Thellefsen, T., Sørensen, B. and Thellefsen, M. (2014), "The information concept of nicholas belkin revisited – some semeiotic comments", *Journal of Documentation*, Vol. 70 No. 1, pp. 74-92.

Thompson, F.B. (1968), "The organization is the information", *American Documentation*, Vol. 19 No. 3, pp. 305-308.

Wellisch, H. (1972), "From information science to informatics: a terminological investigation", *Journal of Librarianship and Information Science*, Vol. 4 No. 3, pp. 157-187.

Yovits, M.C. (1969), "Information science: toward the development of a true scientific discipline", *American Documentation*, Vol. 20 No. 4, pp. 369-376.

Zhang, Y. (1988), "Definitions and sciences of information", *Information Processing & Management*, Vol. 24 No. 4, pp. 479-491.

Zins, C. (2007), "Conceptual approaches for defining data, information, and knowledge", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 4, pp. 479-493.

About the author

Liangzhi Yu is a Professor in the Department of Information Resource Management of the Business School of Nankai University, where she teaches Introduction to Library and Information Science and Library Management. She earned her Doctoral degree from the Loughborough University, UK. She has conducted research on a range of LIS topics; her most recent research focusses on information inequality and public library development in China. Professor Liangzhi Yu can be contacted at: lzhyu@nankai.edu.cn