Combining High Metainformation
with High Information Content:
The Information-Metainformation
Utility Hypothesis

Robert Losee

CB#3360, School of Information and Library Science, UNC-Chapel Hill,
Chapel Hill, NC 27599-3360 <loseeunc@gmail.com>

Robert Losee is a Professor in the School of Information and Library Science at the University of North Carolina at Chapel Hill. His interests are in organizing information, information retrieval, and the study of information and knowledge. His most recent book, *Information From Processes: About the Nature of Information Creation, Use, and Representation*, addresses the nature of information and knowledge, providing a precise definition for both. He has a strong interest in both Library Science and Information Science and tries to bring the two together whenever possible.

Abstract: Many documents and other informational objects carry both information and metainformation about the original informational object. There are general characteristics for documents or objects that possess either high levels of information and high levels of metainformation, or high levels of information and low levels of metainformation, or low levels of information and high amounts of metainformation, or low amounts of information and low amounts of metainformation. Each of these combinations represents a frequently occurring type of informative object. We suggest an Information-Metainformation Utility hypothesis that predicts that the expected economic value of information and its associated metainformation is proportional to the combined amounts of information and metainformation. The use of rules consistent with this hypothesis is discussed. This may be applied to any situation where there is either information or metainformation that may or may not be acquired or used, along with the expected value of the informative object. The idea of ideological segregation, where people tend to view media that represents their prior political beliefs, is examined in this context.

1.0 Introduction

The relationship between the amount of information in a document, or any other type of informational object, and the amount of metainformation available, that is, the metadata that describes the informational object, is key for information professionals, especially those associated with the design and application of metainformation systems through indexing and cataloging in the library and information and knowledge professionals (Gnoli 2012; Losee 1993; Smiraglia & van den Heuvel 2013). Understanding the usefulness of different combinations of in-

formation and metainformation is the focus of the discussion below.

The quantity of both metainformation and information itself can vary, with some statements containing greater or lesser amounts of information and, separately, greater or lesser amounts of metainformation (Losee 2012). The information is referred to here as in a statement, a representation of characteristics of the object being represented. This informative statement is said to be about the characteristics and the entity having these characteristics. A statement may be seen as the representation of the characteristics, and a statement may be understood

124

Knowl. Org. 41(2014)No.2

R. Losee. Combining High Metainformation with High Information Content

as the information in the set of characteristics being represented. Information may take a number of forms, and has been defined many different ways (Gleick 2011; Losee 1997).

The presence of higher levels of metainformation and information together often has a greater expected economic value to the recipient of the informative object than other levels of information and metainformation. As an example, books or webpages that are the most popular and thus most valuable are those that have information not presented elsewhere and that have the most metainformation about them, such as book reviews, hyperlinks to web pages, and so forth.

Knowledge may be viewed as the study of the organization of concepts, using the concepts' characteristics and their arrangement (Dahlberg 2006). However, knowledge also may be rigorously defined as a combination of various forms of information (Losee 2012, pp. 167-179) and knowledge is largely viewed this way below. For example, metainformation may be combined with information to produce knowledge. When an informative statement is true, believed, and justified, it may be said to be a known statement (Audi 2011; Bostock 1988; Gettier 1963; Losee 2012; Lycan 2006). The truth of a statement is an attribute of a statement that exists or does not exist. The belief and the justification for a statement often exist as metainformation about an informative statement, and combined, their presence may transform an informative statement into a known statement. The value placed on knowledge in our culture may serve to reward people who bring together useful information and associated metainformation that, combined, transforms informative statements into known statements.

1.0 Literature

Knowledge organization may be understood as the arranging of concepts, based in part on the characteristics of the concepts (Dahlberg 2006). This approach to the understanding of knowledge allows one to view both knowledge and its organization as separate components which, when combined, produces knowledge organization. Knowledge may also be defined as the relationship between informative statements, combined quantitatively or qualitatively or both (Li, et al. 2004; Losee 2012), while the relationships between objects or concepts may be studied by examining the quantitative and qualitative relationships between these objects or concepts, often in a geometric context (Losee 1992).

Metainformation is information about the information at the output of a process. Often containing characteristics about the informative process, metainformation factors may include who, what, when, where, and why, with many other taxonomies and classes of types and meanings being used in different systems (Greenberg 2009; Willis, et al. 2012; Smiraglia 2005). Information professionals have developed and used standards for the application of metainformation through the use of ontologies, controlled vocabularies in thesauri, and cataloging standards. These support the arrangement of concepts and informative statements, producing knowledge organization.

Studies have shown the utility of indexing and metadata or metainformative terms in the retrieval of documents (Cleverdon 1967; Salton & Lesk 1968). Academic libraries receive many requests for material based on author and title information, but many other requests are subject requests. The Cornell University Libraries found that 61% of searches there targeted authors and titles, with many others being subject searches (Banush & LeBlanc 2007). Some subject requests can be satisfied by matching query terms with terms in document titles, but many subject queries do not match with any title terms, requiring some indexing of documents for those documents to be retrieved, given subject term queries. For example, Gross and Taylor (2005) found that about 1/3 of the records they examined that did not have subject headings would not be retrieved by keyword searching.

The application of classification systems provides a form of metainformation about an informative object. Some classification systems are broad and general, while others provide a great deal of information about each object. A book might be labeled as being solely about history, or further information might be provided that it is about French history. The *International Classification of Disease* has gone through multiple editions, with the latest edition, ICD-11, being more detailed than its predecessors, with the average information per item being greater than the average amount of information found in previous editions, as evidenced by the number of categories and the increased semantic specificity (Bar-Hillel & Carnap 1953).

Metainformation can take a number of forms that exist within the field of bibliometrics. Citations in academic papers indicate materials that were useful in the research described in the article. A hyperlink from one web page to another indicates that the web page that contains the hyperlink contains metainformation about the web page being pointed to. Citations to material and hyperlinks to material on the Internet are indicators of referencing. These citations and hyperlinks sometimes indicate popularity or usefulness.

Information is the content of the output of a process. As a process accepts any available inputs and manipulates it, producing the output, the output can be said to be *about* the input and the processing itself. Information and

metainformation may be examined both qualitatively and quantitatively. One can examine the content of an informative statement. One can also use any of a number of measures of information to compare the amount of information in two different informative statements.

Differing statements may be made about an informative object. For example, different people see different things when observing an object. Those with professional expertise in an area observe the subject of their discipline differently than do lay people, even when observing the same phenomena. When we refer to informative statements and the amount of information present, the amount of information present is that amount of information received by the observer. An encyclopedic work may contain large quantities of information, but from the standpoint of a user, all that is relevant may be the information in a single article. Similarly, a massive database may contain extremely large quantities of information, but the amount of information that the observer receives may be a small amount retrieved from entering a single query to the database system.

Information may be defined using common notions. For example, the character Cookie Monster on Sesame Street defined information as "news or facts about something" (Losee 1997). Academics usually provide more rigorous definitions of information that apply primarily to their own academic field. Here we use a more general definition that is rigorous but can be applied to all academic disciplines. Similarly, our notion of metainformation is broad and universally applicable. We do not claim that this is the only approach to information that is both rigorous and generally applicable, but we do find it useful as a tool and hope that others arguing for different definitions of information can apply their models as generally and as rigorously. We desire to use a rigorous but general definition of both information, in the form of language, bits, objects, or representations, as well as a rigorous but general definition of information that can be applied to metainformation that is, in turn, about language, bits, objects, or representations.

The relative economic value of information and its quantification may be studied in different ways (Rasmusen 2006; Losee 2012). For example, Saracevic and Kantor (1988) provided economic data associated with users' perceptions of utility for documents, providing a basis for estimating the value of assigning metadata to the documents (Cooper 1978). The value of metainformation supplied by systems can be measured for organizational purposes, including evaluation, accounting, and workflow analysis (Greenberg, et al. 2013). Generally, the value of information to a decision maker takes the form of additional information about the true state of nature, reducing uncertainty and increasing the expected economic value of deci-

sions made because of more knowledge about the true state of the world. Art and fiction have a different kind of value, with the information itself having an intrinsic value because of what it does for us and to us, due to its impact upon the individual's neural and emotional states. Rational models of decision making provide specific models for valuing information (Rasmusen 2006), while the value of information that appeals to our subjective selves is based upon more subjective factors, although these informational objects may still be ordered by preference and thus serve as measures.

Informative statements may be believed due to trust that is held with a high degree of confidence, so that one can state, for example, *I trust that the sun will set this evening*. One can speak of person X trusting person Y when X believes that it is highly likely that anything that Y says to X is true. Trusting that information is correct, or lacking the trust that it is correct, represent important forms of metainformation: belief that something is probably true because of who spoke the informative statement, or, in other words, a belief in information because of the associated metainformation. Trust, as a kind of metainformation about an informational object, can be increased or decreased, often through behaviors. By repeatedly supporting someone or being truthful to that person, trust is enhanced and one is more likely to believe a statement provided by that trusted person than when the statement comes from an untrusted person (Fukuyama 1996). Metainformation, as perceived by the potential user, is associated with how much the user trusts the metainformation, and this trust in the metainformation is associated with a degree of belief in the metainformation. A trustworthy source of information about the informative object might include being told by a trusted friend or close family member, or an informative statement in a trusted library reference. Contracts can be developed so that an incentive is provided when a party acts in a trustworthy manner, ensuring that it is to the metainformation provider's advantage to provide accurate metainformation. Reputation is a result of repeated actions that develop a trust in statements accuracy; one trusts material in a journal in part because of the reputation of the journal, developed by the journal through repeatedly publishing material that is found to be accurate.

Below, we speak loosely about relatively high and relatively low information statements, based on the quantity of the information in the statement, as well as relatively high and relatively low metainformation, depending on the amount of information or metainformation in a statement. These informative statements serve as components of statements of knowledge. Examining these kinds of statements may allow for some general principles about quantities of information and metainforma-

126

Knowl. Org. 41(2014)No.2
R. Losee. Combining High Metainformation with High Information Content

tion to be considered and evaluated. A highly informative statement contains more information than a low information statement. We do not rigorously define "high" or "low" amounts of information or metainformation, instead treating them as relative terms, without specifying the cutoff between high and low. The amount of information in an informative statement here can be understood as inversely related to the probability of the statement; these statements can be studied as-is or the reduction in uncertainty due to the receipt of a statement can be measured (Gleick 2011; Losee 2012; Shannon 1993). A statement with a great quantity of information can be said to have high content or high information. Adding material to a statement will often make it more informative (Bar-Hillel & Carnap 1953). Similarly, metainformation may be determined to be more or less informative than another metainformative statement, by studying the amount of information in the metainformation. The amount of knowledge is dependent on the amount of information and the amount of metainformation in the statements that lead to a known statement. Metainformation is usually present when informative statements are known statements.

2.0 High Metainformation and High Information

Statements that are relatively high in both amounts of information and metainformation serve as the most economically beneficial statements, on the average, when compared to statements with relatively low quantities of information and low quantities of metainformation. There are commonalities between the natures of types of informational objects and how we relate to them, in terms of whether they are highly informational or low informational, or whether they have a relatively large amount of metainformation or a relatively small amount.

Consider a recently published book. If it is a bestseller, it will be somewhat unique and thus have more information in some senses; a popular self-help book provides new ideas, while a new romantic novel provides some special plot point that makes it more attractive. The metainformation available to end users includes a range of comments and reviews, as well as numerous advertisements; all of these being information about the book, and that are also associated with the book's expected worth to readers.

Other informative objects also serve as most useful when they are very informative and have greater amounts of metainformation. Material placed in libraries and archives is usually more informative than non-acquired materials and possess more metainformation than average. Having more information, with unique aspects, makes the material more useful for library patrons than other less informative materials. Additionally, the kind of material that is acquired by libraries often tends to be more popular and contains more metainformation than non-acquired materials. Some of this metainformation is provided by librarians in the form of bibliographic records and other information assigned to a document, while material that has more reviews is more likely to be acquired by libraries. A book by a Nobel Prize winning author is probably unique and represents a cutting edge perspective or a uniquely deep presentation, probably containing more information than an average book. Similarly, it has probably received more book reviews and more favorable book reviews than most other books, because of its expected quality and uniqueness.

Given the meeting of two political leaders in a public place, there will be a high degree of importance attached to their meeting, which may be more widely reported than meetings between "commoners," such as two students in an organization of information class. Those political leaders meeting are likely to say more profound things than are usually uttered when two people meet for the first time. Political leaders often provide more information to each other and they provide more information to the observers. There is more metainformation available about the individuals; for example, there likely is a Web page about each of the famous greeters.

The combination of high information and high metainformation produces a greater economic benefit than other combinations described below. The most beneficial informational objects have the most information, and they are validated as worthwhile by the presence of the most metainformation.

3.0 Low Metainformation and Low Information

Low information occurs when there is a very uninteresting, highly redundant informative object and the informative object is common and there is little metainformation. Algorithmic information theory (Chaitin 1987) suggests that the amount of information in an object is proportional to the size of the smallest program that can produce the original input. A sequence of one million "1" characters might be produced by a computer program containing one million print statements, each producing a single "1," or by a smaller program containing a loop that one million times executes a single statement that prints out a single "1," thus producing one million "1"s. This latter program is much shorter than one with one million individual print statements, with this smaller program showing the relatively small amount of information in this program, whether the program that produces the "1"s is written in this small, loop-based, form, or in the larger program with the one million print statements.

When there is a regularity in the informative object, there is much less information than when there is a great deal of irregularity. A description of a white ceiling might be a very small description, making it a carrier of little information. Even if there is an irregularity in the view of the ceiling, such as a fly on the ceiling, the description will be relatively small when compared to the description of a ceiling with numerous randomly positioned spots on it, which will have precise descriptions of the size, shape, color, and location of each of these spots. Related to this is the idea that adding specificity to a statement decreases the number of statements with which it is consistent and increases the information in the statement (Bar-Hillel & Carnap 1953).

A description that is relatively low in information and also low in metainformation about the informational object is generally of little utility, compared to a high information-high metainformational object. An untrusted description of a plain ceiling is of little benefit, while an accurate description of a beautiful painting, that captures all the nuances, perhaps described by an art critic that points out many relationships in the painting that one wouldn't notice normally, would be highly appreciated. The provider of some information might lie about what is being represented, producing information that is of little expected value. Repeated lies and the consequent lack of trust might often produce low metainformation and often low information, with overall low expected economic benefit to a decision maker. Information with a high amount of redundancy, such as common patterns like breathing, may have little overall information and relatively little metainformation. Generally such patterned information is of little overall benefit.

4.0 Low Metainformation and High Information

Everyday objects that we routinely see around us often have unique and sophisticated characteristics but have little metainformation. We do not know their origin or how they function. Yet, for our purposes, they perform in unique and often complex ways that benefit us. Low metainformation may occur because one cannot trust the information or because we know little or there is little easily available about the informative object. A young child viewing a bird knows virtually nothing about birds other than that they fly and sing. The provider of information may be a frequent liar or often inaccurate and thus provides little accurate metainformation. Note that we treat metainformation as accurate metainformation that reflects the true state of the world; information or metainformation that is false is not about the informative object, but instead represents information or metainformation about some other informative object, such as a liar, instead of what the liar describes. In

libraries, some material may be informative but not worth the effort of applying much metainformation, in some people's opinions. For example, some institutions may use "quick cataloging," a brief form of cataloging that provides less metainformation than full cataloging does, for popular new fiction that is expected to be kept in the collection for only a relatively brief period of time or will not receive much use.

Certain natural events, such as the radioactive decaying of uranium molecules, occur in a random manner that provides a great deal of information, in that it isn't predictable. However, there is almost no metainformation about this process. Computer algorithms can generate pseudo-random numbers that can be used in place of true random numbers. These also have a low amount of metainformation but high amounts of information. Note that when the information in these situations is high and the informative objective is valuable, the metainformation is usually extensive.

5.0 High Metainformation and Low Information

High metainformation and low information occur together in situations where there is a great deal of context, but the actual information itself is minimal or largely irrelevant. For example, amateur radio operators often try to contact other countries. In contests, they don't wish to talk much with the person on the other end of the connection; they want the thrill of talking with many people far away, that is, the contact itself. Here the metainformation, such as the country contacted or the time of day, is very important, but the information exchanged in the contact is often not important. In contests, "ham" operators will often transmit a signal report, which is almost always "59" in voice or "599" in Morse code, meaning that the signal was received clearly, along with a small piece of information, such as what state they are in, if in the United States, what their age is, or a contact number. Similarly, shaking hands with a major politician or actor is long remembered, but little information is exchanged in the information interchanging process itself.

A highly contextualized note with only the word "bread" on it in my spouse's handwriting probably represents something I need to purchase at the store. I know by the style of the note and its contents that it is probably a request for an action. Such a simple note has an important context or metadata, and is a common example of an informative object that has little information but quite a bit of metainformation (including what will happen if I don't purchase bread, where I might purchase bread, etc.). Very small documents in libraries and archives may say very little directly but have much metadata and context produced about the original information. A

128

Knowl. Org. 41(2014)No.2
R. Losee. Combining High Metainformation with High Information Content

letter from a president to someone else may be directly inconsequential, but many archives would describe many characteristics of the letter so that it might be easily located, used, appreciated, and understood.

6.0 Combining Information and Metainformation to Produce Economic Value

The content of an informational object, along with its metadata, may be used together by a decision maker to produce an action that would have an expected economic value, often because the information is transformed into knowledge (Losee 2012). The joint increases of the amount of information and metainformation increases the likelihood of increased knowledge. This economic value reflects the worth of the action the decision maker takes or might take, given a set of repeated situations. For example, journals that aren't indexed and have less metainformation are worth much less than journals that are indexed with more metainformation, everything else being equal. Similarly, journals with a great deal of information are worth much more than otherwise equal journals that have little information.

We propose the following:

> **Information-Metainformation Utility Hypothesis**: On the average, the greater the combined amount of information and metainformation, the greater will be the expected utility of the informational object. Conversely, the lower the combination of the amount of information and metainformation, the lower will be the expected utility of the informational object.

The nature of the combination operation needs to be such that the combination of two positive values is always greater than either one of the individual values. One might think of combining amounts using arithmetic operations such as addition or multiplication. Both of these operations produce a value greater than either of the individual values. The exact nature of the combining operation remains to be determined: it is likely to be different for different kinds of situations, and it may not be commutative. Lack of commutativity means that 3 units of information and 4 of metainformation wouldn't necessarily produce the same result as 4 units of information and 3 of metainformation.

Kenneth Haase (2004) suggests that "As media technologies improve and spread, there will be a gradual transformation where metadata will become more valuable (on average) than the content it describes." Haase refers to this as a *metadata twist*. If the Information-Metainformation Utility hypothesis is true, and the

amount of information remains constant, the expected utility of information combined with metainformation, often combined to be knowledge, is likely to increase over time.

7.0 An Indexing and Acquisition Rule

A model relating variables is always useful when using or practicing science, and a model combining information and metainformation, producing expected economic value, can be applicable in a range of areas. The combination operator "*" is used below, with the asterisk representing a combination of two amounts. The combination

*Information * Metainformation* is proportional to *Utility*

formalizes that the amount of information combined with the amount of metainformation provides a value that is proportional to the expected utility of possessing the combination (see Figure 1.)

		Information	
		Low	High
MetaInformation	Low	Lowest Expected Economic Value	
	High		Highest Expected Economic Value

Figure 1. The relationships between Information, Metainformation, and the Expected value of informative objects.

Information professionals may use this relationship in deciding when to assign metadata to documents or informative objects. One can compute raw values, such as the amount of information in a book to be cataloged, the expected value of the assigned metainformation, as well as the expected value of the book. A second approach, probably easier to apply, is to compare a document being assigned an index term with a gold standard document with a known and established information amount, metainformation amount, and an expected economic value. By comparing the object of interest to the gold standard with regards to some of the established features, a limited range of values can be estimated for the other features in the object of interest (Cooper 1978).

The expected utility might represent either a specific value associated with a specific informative object or the expected utility may represent the average of the utilities over a set of informative objects. We might choose to acquire an informational object if the expected amount of information in the object combined with the amount of metainformation exceeds the expected utility. There may

be a tradeoff, with a decrease in information being compensated for by the presence of more metainformation.

By considering the rules above and specific numeric values, one could estimate whether an informational object and the associated metainformation are high information—high metainformation or whether they are low information—low metainformation, or one of the other two possible "high-low" categories. By having specific cutoffs for the categories, one might be able to label an informative statement as falling into one of the four categories. Believing that a statement is in one of the four categories and having two of the values from either the amount of the information, metainformation, or expected economic value, one can estimate the third, remaining value. There are significant applications of this hypothesis to human decision making.

8.0 Media and Information Sources

Most individuals value some media while finding other media worthless, occasionally malicious (Gentzkow & Shapiro 2011). Media consumers often engage in ideological *segregation*, the choice to be with others of similar interests. One is most likely to read material that one agrees with rather than material with which one would disagree. Clearly the value of a given media source is based upon our own tastes. As was suggested earlier, humans value items such as media based upon both the amount of information it provides and the amount of metainformation that we have about the information.

Information is present in all media, but much of it is not absorbed when reading or viewing. Reading material from a source that one doesn't like often involves skimming the material or reading it in more depth with a predisposition to viewing it as wrong or carrying incorrect ideas. The information that is absorbed into the brain through observation is thus tainted. The amount of information absorbed is far lower if the media represents opposing viewpoints than if the informative object represents a favorable position.

The amount of metainformation present in a source we like is much greater than for a source we oppose. We possess metainformation ourselves about those things with which we are familiar, allowing us to add to the metainformation that arrives from the outside world. While there are some individuals who practice the "know your enemy" approach to gathering information and choose to know everything possible about opposing views, most people read primarily about topics and issues they find interesting. The metainformation about these authors and issues increases as one reads more.

The information and metainformation that we choose to absorb is that which has the greatest expected economic value. This value is related to the amount of information we absorb from the informative object, along with the metainformation. Because we retain some metainformation and some information, those informational objects that are consistent with what we retain are expected to have the highest economic value to information consumers. The ideological segregation is produced through a feedback cycle. The more one consumes about a topic, the expected amount of information that one absorbs and the metainformation that one possesses increases. The expected value of these increases, as one increasingly reads from one's favorite media outlets.

9.0 Conclusions

Metainformation, as well as information, can be present in high amounts or low amounts, often producing knowledge. The combinations of low and high information and metainformation have unique characteristics that may be useful for information professionals making decisions. Having these categories helps us to make practical decisions by considering the type of an informative object. The combinations of different levels of information and metainformation are hypothesized to have economic value to the recipient of the information that is proportional to the combined amounts of information and metainformation. Using this economic relationship between information and metainformation helps us to better describe, predict, and understand both metainformation and information itself in a human and economic context, providing a core relationship in the study of indexing, metadata, and knowledge organization. Examples of these relationships are given; further studies may examine more quantitatively these economic relationships between information and metainformation. Information professionals can use such rules by determining both the potential value of assigning metadata, as well as determining the potential utility of informative objects when metainformation is assigned, sometimes producing knowledge.

References

Audi, Robert, 2011. *Epistemology*. Third ed. New York, NY: Routledge.

Banush, David and LeBlanc, Jim. 2007. Utility, library priorities, and cataloging policies. *Library collections, acquisitions, and technical services* 31: 96-109.

Bar-Hillel, Yehoshua and Carnap, Rupert. 1953. Semantic information. *The British journal for the philosophy of science* 4: 147-57.

Bostock, David, 1988. *Plato's theaetetus*. Oxford: Clarendon.

130

Knowl. Org. 41(2014)No.2

R. Losee. Combining High Metainformation with High Information Content

Chaitin, Gregory J., 1987. *Algorithmic information theory*. Cambridge, England: Cambridge University Press.

Cleverdon, Cyril. 1967. The Cranfield tests on index language devices. *ASLIB Proceedings* 19: 172-81.

Cooper, William S., 1978. Indexing documents by Gedanken experimentation. *JJournal of the American Society for Information Science* 29: 107-19.

Dahlberg, Ingetraut, 2006. Knowledge organization: a new science? *Knowledge organization* 33: 11-9.

Fukuyama, Francis, 1996. *Trust: the social virtues and the creation of prosperity*. New York: Free Press Paperbacks.

Gentzkow, Matthew and Shapiro, Jesse M., 2011. Ideological segregation online and offline. *Quarterly journal of economics* 126: 1799-839.

Gettier, Edmund L. 1963. Is justified true belief knowledge? *Analysis* 23: 121-3.

Gleick, James. 2011. *The information: a history, a theory, a flood*. New York: Pantheon.

Gnoli, Claudio. 2012. Metadata about what? distinguishing between ontic, epistemic, and documental dimensions in knowledge organization. *Knowledge organization* 39: 268-75.

Greenberg, Jane. 2009. Metadata and digital information. In: *Encyclopedia of library and information science*. New York: Marcel Dekker, pp. 3610-23.

Greenberg, Jane, Hodge, Gail, Szostak, Rick and Uckun, Timur Han, 2013. Net gain via knowledge organization: classification and productivity. Presented at *ASIST 2013, November 1-6, Montreal, Canada*. Available http://www.asis.org/asist2013/proceedings/submissions/panels/43panel.pdf.

Gross, Tina and Taylor, Arlene G., 2005. What have we got to lose? the effects of controlled vocabulary on keywords searching results. *College & research libraries* 66: 212-30.

Haase, Kenneth. 2004. *Context for semantic metadata*. In *Proceedings of the 12th annual ACM international conference on Multimedia*. New York: ACM, pp. 204-11.

Li, Ming, Chen, Xin, Li, Xin, Ma, Bin and Vitanyi, Paul M.B. 2004. The similarity metric. *IEEE transactions on information theory* 50: 3250-64.

Losee, Robert M. 1992. A Gray code based ordering for documents on shelves: classification for browsing and retrieval. *Journal of the American Society for Information Science* 43: 312-22.

Losee, Robert M. 1993. Seven fundamental questions for the science of library classification. *Knowledge organization* 20: 65-70.

Losee, Robert M. 1997. A discipline independent definition of information. *Journal of the American Society for Information Science* 48: 254-69.

Losee, Robert M. 2012. *Information from processes: about the nature of information creation, use, and representation*. Berlin: Springer Verlag.

Lycan, William G., 2006. On the Gettier Problem problem. In Hetherington, Stephen Cade, ed., *Epistemology Futures*. Oxford: Oxford U. Press, pp 148-68.

Rasmusen, Eric. 2006. *Games and information*. 4th ed. Oxford: Wiley-Blackwell.

Salton, Gerard and Lesk, Michael. 1968. Computer evaluation of indexing and text processing. *Journal of the ACM* 15: 8-36.

Saracevic, Tefko and Kantor, Paul. 1988. A study of information seeking and retrieving. II. Users, questions, and effectiveness. *Journal of the American Society for Information Science* 39: 177-96.

Shannon, Claude Elwood, 1993. A mathematical theory of communication. In his, *Claude Elwood Shannon: collected papers*. New York: IEEE Press, pp. 5-93.

Smiraglia, Richard P., 2005. *Metadata: a cataloger's primer*. Binghamton, NY: Haworth Information Press.

Smiraglia, Richard P. & van den Heuvel, Charles, 2013. Classification and concepts: towards an elementary theory of knowledge interaction. *Journal of documentation* 69: 360-83.

Willis, Craig, Greenberg, Jane and White, Hollie C. 2012. Analysis and synthesis of metadata goals for scientific data. *Journal of the American Society for Information Science and Technology* 63: 1505-20.